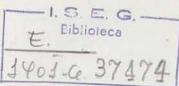




RESERVADO



HA31.3. P76 1990

MÉTODO "BOOTSTRAP"

Uma Aplicação na Estimação e Previsão
em Modelos Dinâmicos

Isabel Maria Dias Proença

Dissertação apresentada como requisito parcial para a obtenção do grau de
Mestre em Métodos Matemáticos para Economia e Gestão



Ao Sr. Professor Doutor Bento Murteira, orientador desta dissertação, é devido um profundo agradecimento pelo auxílio prestado em todos os momentos da sua realização.

Agradeço ainda a todos os colegas e amigos que de algum modo contribuíram para a realização deste trabalho, muito em especial a Isabel Andrade e a João C. Santos Silva.



1 - INTRODUÇÃO	1
2 - METODOLOGIA <i>BOOTSTRAP</i>	9
2.1 - Descrição do <i>bootstrap</i>	9
2.2 - Reamostragens alternativas	13
2.3 - Propriedades assintóticas do <i>bootstrap</i>	20
3 - INTERVALOS DE CONFIANÇA <i>BOOTSTRAP</i>	32
3.1 - Método dos percentis	37
3.2 - Método com correcção do enviesamento	46
3.3 - Método com correcção do enviesamento e aceleração	53
3.4 - <i>Bootstrap-t</i>	80
3.5 - Desenvolvimentos recentes	82
3.6 - Algumas considerações	86
4 - O <i>BOOTSTRAP</i> NA REGRESSÃO	91
4.1 - Caracterização do modelo	91
4.2 - Algoritmo <i>bootstrap</i>	93
4.3 - Propriedades assintóticas	96
4.4 - Algumas considerações	100

5 - APLICAÇÃO EMPÍRICA: ESTIMAÇÃO E PREVISÃO DO CONSUMO	104
5.1 - O <i>bootstrap</i> na regressão	105
5.2 - Previsão do consumo	126
6 - CONSIDERAÇÕES FINAIS	137
ANEXO 1 - TESTES DE ESPECIFICAÇÃO DO MODELO	A-1
ANEXO 2 - MODELIZAÇÃO DE DY	A-10
BIBLIOGRAFIA	B-1

ABREVIATURAS:

- C.F. - Cornish-Fisher;
- d.p. - Desvio padrão;
- E.E. - Expansão de Edgeworth;
- I.C. - Intervalo de confiança;
- M.V. - Máxima Verosimilhança;
- Φ - Função de distribuição $\mathcal{N}(0,1)$;
- ϕ - Função de densidade de probabilidade da $\mathcal{N}(0,1)$;
- z_α : $\Phi(z_\alpha) = \alpha$;
- $F^{-1}(\alpha) = \sup \{x: F(x) \leq \alpha\}$;

1 - INTRODUÇÃO

O *bootstrap* foi introduzido por Efron(1979) como metodologia geral para avaliar estatisticamente a qualidade de um estimador. Insere-se num conjunto de técnicas (essencialmente não paramétricas) de reamostragem; é flexível e de largo alcance e baseia-se no princípio da Máxima Verosimilhança (M.V.). Evita a análise teórica dedutiva, substituindo-a pela computação intensiva. Descarta a modelização, a necessidade de suposições ou de hipóteses e pode portanto aplicar-se de forma automática e simples, a um vasto conjunto de situações por mais complexas que sejam.

A expansão deste tipo de métodos foi facilitada pelo desenvolvimento da informática que trouxe uma crescente velocidade e acessibilidade do cálculo computacional. Como Efron refere¹,

The remarks so far seem to indicate a statistical world in which mathematical theory is replaced by brute computation. In fact, computational advances precipitate new theory.

Este trabalho pretende fazer uma abordagem global da teoria *bootstrap*, expondo a sua lógica na determinação de estimativas do enviesamento, desvio padrão, e distribuição empírica de uma estatística, assim como na construção de I.C. aproximados, apontando as respectivas propriedades assintóticas.

Simultaneamente procurou fazer-se uma aplicação das técnicas expostas. Dos inúmeros campos onde a aplicação do *bootstrap* pode ser proveitosa, seleccionou-se a regressão.

¹ Efron(1979) p.174.

O capítulo 2 apresenta a metodologia de reamostragem *bootstrap* para um problema geral não paramétrico, tratado nos textos iniciais de Efron. São expostas as estimativas *bootstrap* do enviesamento e do desvio padrão de uma estatística assim como a determinação da distribuição amostral aproximada.

A ideia chave do *bootstrap* consiste em transformar o problema não paramétrico inicial num problema paramétrico artificial onde a distribuição desconhecida da população é substituída pela distribuição empírica da amostra, sendo restringido universo das v.a. aos valores observados. Desta forma as estimativas de $E_F\{\theta\}$ e $V_F\{\theta\}$ são respectivamente $E_F\{\hat{\theta}\}$ e $V_F\{\hat{\theta}\}$, enquanto que $P_F\{\theta \leq k\}$ tem como estimativa $P_F\{\hat{\theta} \leq k\}$. Estes resultados coincidem com as estimativas tradicionais da M.V. não paramétrica. No entanto, em muitas situações, as expressões analíticas das quantidades a estimar não são conhecidas ou, por outro lado, as respectivas estimativas são demasiado complexas de deduzir, invalidando a aplicação dos métodos clássicos. É nestes casos que o *bootstrap* revela uma supremacia sobre a teoria tradicional ao possibilitar a obtenção de soluções numéricas através do algoritmo de simulação tipo Monte Carlo.

Referem-se também outras formas de reamostragem *bootstrap* nomeadamente o *bootstrap* alisado e o *bootstrap* paramétrico. Esta última permite a generalização aos problemas não paramétricos, pois como refere Efron,¹

Fisher's familiar theory for assigning a standard error to a maximum likelihood estimate is itself a "bootstrap theory" carried out in a parametric framework.

Ainda no capítulo 2 analisam-se as propriedades assintóticas do *bootstrap* para um conjunto de estatísticas. Inicialmente pretende mostrar-se a validade assintótica provando que conduz aos limites assintóticos clássicos conhecidos. Este resultado

1 Efron(1982a) p.29.

contudo não realça as potencialidades do *bootstrap* relativamente aos métodos clássicos. No entanto, o conhecimento das propriedades assintóticas do *bootstrap* foi enriquecido com a dedução por alguns autores de taxas de convergência das estimativas *bootstrap*. Estes resultados recorrem a expansões de Edgeworth que consistem em desenvolver a distribuição de uma v.a. numa série de vários termos em que o primeiro é uma Normal e os outros envolvem sucessivas potências de $1/\sqrt{n}$ (com expoente crescente).

Estas taxas deixam patente que, em muitos problemas, a convergência da aproximação *bootstrap* é mais rápida do que a convergência à Normal. Noutra perspectiva, é possível concluir que a grandeza do erro da aproximação *bootstrap* é menor do que a grandeza do erro decorrente da aproximação Normal. Autores referem a propriedade do *bootstrap* de "correção parcial" do termo $n^{-1/2}$, isto é, o *bootstrap* aproxima correctamente parte do segundo termo (termo em $n^{-1/2}$) da expansão de Edgeworth da distribuição exacta.

Ainda sobre a superioridade do *bootstrap* relativamente à aproximação Normal, Beran recorrendo à teoria da decisão estatística, prova que a distribuição *bootstrap* é assintoticamente *minimax* propriedade que se não verifica para a Normal.

Por outro lado, verifica-se que as taxas de convergência *bootstrap* são mais rápidas para estatísticas *studentizadas*. Neste caso, existe mesmo um fenómeno de "correção total" do termo em $n^{-1/2}$.

No capítulo 3 são expostos os métodos de determinação de intervalos de confiança (I.C.) *bootstrap*: percentis, correção do enviesamento (BC), correção do enviesamento e aceleração (BC_a), e *bootstrap*-t. São ainda referidos alguns desenvolvimentos recentes verificados neste campo.

A determinação de I.C. constitui um dos temas fulcrais da teoria *bootstrap*, tema que, longe de estar encerrado, demonstra grande vitalidade surgindo constantemente novas abordagens. Daí

que a literatura sobre este assunto seja denotadamente vasta, donde se explica que o capítulo 3 seja mais extenso que os restantes.

O teorema do limite central, estabelecendo em muitos problemas a convergência para a distribuição Normal, constitui a base teórica dos I.C. aproximados clássicos (intervalos *standard*). Estes intervalos são muito úteis dada a sua vasta aplicabilidade de uma forma simples e automática. No entanto, podem não ser muito rigorosos, nomeadamente por não conseguirem captar a assimetria dos verdadeiros intervalos. Para aumentar a eficácia dos intervalos *standard* criou-se um conjunto de "artifícios" com o intuito de acelerar a convergência para a Normal, como por exemplo a aplicação de transformações prévias às estatísticas como a conhecida transformação do coeficiente de correlação devida a Fisher.

Os I.C. de Efron têm subjacente a ideia da aplicação deste tipo de transformações. Estipulam a existência (pelo menos aproximada) de uma transformação normalizadora embora não seja necessário o seu conhecimento para a determinação dos intervalos. O método dos percentis, que consiste em considerar como extremos do I.C. os percentis da distribuição *bootstrap* da estatística, admite que existe uma transformação da estatística que conduz a uma variável aleatória com distribuição Normal de média igual ao transformado do valor da estatística na amostra observada e variância fixa. O BC e o BC_a surgem da generalização dos pressupostos subjacentes a esta transformação normalizadora. O BC prevê a hipótese da estatística ser enviesada em termos da mediana, alterando a média da Normal com uma constante que traduz o enviesamento (correção do enviesamento, z_0). O BC_a para além desta hipótese, considera ainda que a variância da transformação não é constante (dependendo do valor do parâmetro), conduzindo à correção da aceleração, a .

A consequência da generalização dos pressupostos na determinação dos I.C. consiste apenas numa modificação da ordem dos percentis a considerar como extremos do I.C., sendo a nova ordem determinada, para o BC em função da correcção do enviesamento, e da correcção do enviesamento e aceleração para o BCa, através de fórmulas muito simples. O cálculo de z_0 é imediato. A determinação de a envolve alguma polémica. Destacam-se duas vias - a de Efron e de Peter Hall. A primeira, resulta na estimativa de uma aproximação desta constante, com expressão simples em problemas não paramétricos. A segunda, deduz a de forma a proporcionar propriedades superiores na cobertura do intervalo BCa (unilateral), embora se limite a um tipo muito particular de problemas, podendo ser estimada de forma operacional.

Os intervalos de Efron para além de se basearem em pressupostos mais gerais do que os *standard* aplicam-se ainda de forma automática e simples. Como referem Efron e Tibshirani¹ ,

... the progression from the standard intervals to the BCa method is based on a series of increasingly less restrictive assumptions... Each successive method... requires the statistician to do a greater amount of computation; first the bootstrap distribution, then the bias-correction constant, and finally the constant a. However, all of these computations are algorithmic in character, and can be carried out in an automatic fashion.

O *bootstrap-t* preconiza a construção de I.C. de uma forma mais operacional na medida que não necessita do cálculo de correcções. A lógica é idêntica à de um intervalo clássico com base na *t* de *student* apenas substituindo os quantis desta distribuição teórica pelos quantis de igual ordem da distribuição

1 Efron e Tibshirani(1986) p.70.

bootstrap da estatística *studentizada*. É de notar que este método só pode ser aplicado quando existe uma estimativa fiável do desvio padrão da estatística.

A análise das propriedades assintóticas destes métodos tem sido feita recorrendo a expansões de Edgeworth (E.E.), e de Cornish-Fisher (inversa da E.E.) que permite o desenvolvimento em série dos quantis de uma dada distribuição. Neste âmbito surgem os conceitos *correcto* de primeira ordem e *correcto* de segunda ordem. Um I.C. é *correcto* até à primeira ordem se aproxima correctamente só até ao primeiro termo da expansão dos extremos do I.C. exacto. A correcção até à segunda ordem traduz-se por uma aproximação correcta até ao segundo termo da referida expansão. Só o BC_a e o *bootstrap-t* conseguem ser correctos até à segunda ordem. No entanto, é possível conjecturar que o *bootstrap-t* apresente uma superioridade na captação das propriedades de terceira ordem, sendo ainda dos métodos analisados aquele que determina na amplitude dos intervalos um erro de grandeza inferior. São também o BC_a e o *bootstrap-t* que apresentam um erro de menor grandeza na (probabilidade de) cobertura dos intervalos unilaterais. É interessante notar que todos os métodos cometem um erro de igual ordem de grandeza na cobertura dos I.C. bilaterais.

Com o intuito de melhorar a cobertura dos I.C. aproximados surgiram recentemente algumas abordagens, referidas no ponto 3.5 deste trabalho, como o *bootstrap-t* modificado, o método de calibragem de Loh e a pré-pivotação de Beran. Têm como contrapartida uma maior exigência computacional e, apesar das suas propriedades não estarem ainda totalmente estudadas, parecem bastante promissores.

No capítulo 4 analisa-se a aplicação do *bootstrap* a problemas de regressão linear múltipla. São focados os problemas em que as variáveis residuais são homoscedásticas e não correlacionadas, admitindo-se ainda a existência de valores desfasados da variável endógena. Para estes casos, é apresentada a expressão particular

de reamostragem *bootstrap* em modelos econométricos sendo ainda abordada a respectiva validade assintótica, com conclusões semelhantes às obtidas no capítulo 2.

Para ilustrar a *performance* das técnicas *bootstrap* escolheu-se o problema da estimação e previsão em modelos dinâmicos autoregressivos. Estes modelos são muito frequentes nos estudos económicos aplicados, embora se manifestem algumas insuficiências no seu tratamento estatístico por parte dos métodos clássicos.

Em modelos com variável endógena desfasada não é possível deduzir as propriedades em amostras finitas do estimador Ordinário dos Mínimos Quadrados (*OMQ*). Sabe-se que não é centrado embora não se conheça o significado do enviesamento, sendo contudo consistente. Em pequenas amostras o estimador *OMQ* não coincide com o estimador da M.V., mas aproxima-se deste com o crescimento da dimensão da amostra. Os resultados clássicos utilizados na inferência do modelo são válidos apenas assintoticamente.

Com o *bootstrap* é possível enriquecer a análise estatística do modelo, podendo obter-se uma estimativa do enviesamento do estimador *OMQ* e a construção de I.C. aproximados. Apesar de não haver nenhum estudo teórico sobre a qualidade da aproximação dos I.C. teóricos com base na t , espera-se que os I.C. *bootstrap* correctos de segunda ordem sejam superiores, nomeadamente na capacidade em captar a assimetria dos verdadeiros intervalos.

Na previsão, a necessidade do *bootstrap* é ainda mais premente. Em modelos dinâmicos, não são válidas as propriedades clássicas que permitem a avaliação estatística das previsões. O erro de previsão tem distribuição desconhecida, mesmo quando as variáveis residuais são Normais. Por outro lado, não existe uma expressão teórica para a variância do erro de previsão. O *bootstrap* ultrapassa esta deficiência na informação teórica, possibilitando a determinação de estimativas da média e variância do erro de previsão assim como I.C. aproximados para as previsões.

Escolheu-se a função consumo como objecto desta aplicação pois é tradicionalmente um modelo com variável endógena desfasada, para além de constituir um tema de relevância para a macroeconomia. Assim no capítulo 5 procede-se à estimação de uma função consumo para o período compreendido entre 1958 e 1985 (inclusivé), determinando-se estimativas *bootstrap* do enviesamento do estimador *OMQ*, e I.C. *bootstrap* para os parâmetros. De seguida, calculam-se previsões das despesas em Consumo para 1986 e 1987, assim como, através das técnicas *bootstrap*, estimativas da média e variância do erro de previsão. São também construídos I.C. para as previsões, apresentando-se as respectivas distribuições empíricas.

2 - A METODOLOGIA *BOOTSTRAP*

2.1 - Descrição do *bootstrap*

Seja \underline{X} uma amostra aleatória de dimensão n , $\underline{X} = (X_1, X_2, \dots, X_n)$, onde as v.a. X_i são independentes e idênticamente distribuídas de acordo com F desconhecida. Para esta amostra observou-se a realização $\underline{x} = (x_1, x_2, \dots, x_n)$. Seja $\theta(F)$ (funcional de F) um parâmetro desconhecido de F e $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ uma dada estatística, estimador de $\theta(F)$, que assume o valor $\hat{\theta}(x_1, x_2, \dots, x_n)$ para a realização da amostra.

O objectivo consiste em determinar com base na realização observada \underline{x} uma estimativa da distribuição de $\hat{\theta}$ ou de algum aspecto particular desta distribuição como, $E_F\{\hat{\theta}\}$, $[V_F\{\hat{\theta}\}]^{1/2}$ ou ainda, $P_F\{\hat{\theta} < k\}$ com, $k \in \mathbb{R}$.

O *bootstrap* responde a esta questão operando uma transformação do problema inicial. Substitui as v.a. X_1, X_2, \dots, X_n pelas v.a. *bootstrap*, X'_1, X'_2, \dots, X'_n , independentes e idênticamente distribuídas, assumindo valores no conjunto $\{x_1, x_2, \dots, x_n\}$ constituído pelos valores da realização \underline{x} . A variável *bootstrap* X'_i ($i=1, 2, \dots, n$) tem distribuição \hat{F} , a distribuição empírica da amostra \underline{X} que confere probabilidade $1/n$ a cada x_i ($i=1, 2, \dots, n$). A amostra \underline{X} é substituída pela amostra *bootstrap* $\underline{X}' = (X'_1, X'_2, \dots, X'_n)$ enquanto que a estatística $\hat{\theta}$ tem como réplica *bootstrap* a estatística $\hat{\theta}' = \hat{\theta}(X'_1, X'_2, \dots, X'_n)$. Desta forma, o problema não paramétrico inicial é transformado num problema paramétrico artificial.

A amostra genérica *bootstrap* $\underline{X}' = (X'_1, X'_2, \dots, X'_n)$ pode caracterizar-se por um vector de reamostragem igual a,

$$\underline{P}' = [P'_1 P'_2 \dots P'_n] \quad , P'_i = \#\{X'_i = x_i\}/n.$$

Este vector tem distribuição multinomial, ponderada pelo factor $1/n$, dado que nP_i^* representa o número de ocorrências da categoria i de entre n categorias possíveis cada uma com probabilidade de realização $1/n$, numa sucessão de n provas independentes,

$$\underline{P}^* \sim (1/n) \text{ Mult}(n, \underline{P}^0) \quad , \quad (2.1)$$

com,

$$\underline{P}^0 = [1/n \ 1/n \dots 1/n]; \ E\{\underline{P}^*\} = \underline{P}^0; \ V\{\underline{P}^*\} = (1/n)^2 I_n - (1/n) \underline{P}^{0T} \underline{P}^0 \quad ,$$

(I_n representa a matriz identidade de ordem n).

A estimativa da distribuição de $\hat{\theta}$ é dada muito simplesmente pela distribuição *bootstrap* (distribuição teórica) de $\hat{\theta}^*$ enquanto que as estimativas de $E_F\{\hat{\theta}\}$, $[V_F\{\hat{\theta}\}]^{1/2}$ e $P_F\{\hat{\theta} < k\}$ são obtidas respectivamente através de $E_F\{\hat{\theta}^*\}$, $[V_F\{\hat{\theta}^*\}]^{1/2}$ ou ainda $P_F\{\hat{\theta}^* < k\}$.

O cálculo teórico directo destes valores só pode ser efectuado em casos muito raros, como no problema em que $\hat{\theta}(X_1, X_2, \dots, X_n)$ é a mediana.¹ Na quase totalidade das situações há que recorrer a aproximações obtidas com a utilização de técnicas de Monte Carlo.

A determinação destas aproximações pressupõe a implementação do seguinte algoritmo *bootstrap*:

Dadas as variáveis aleatórias *bootstrap*, $X_1^*, X_2^*, \dots, X_n^*$, conforme foram definidas,

- 1) Constrói-se uma realização da amostra, seja $\underline{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$, utilizando a estrutura de probabilidade verificada pelas v.a. *bootstrap*. Quando a distribuição de cada $X_i^* (i = 1, 2, \dots, n)$ é \hat{F} , \underline{x}^* resulta de uma concretização do vector de reamostragem \underline{P}^* . Este processo equivale a retirar ao acaso um elemento do conjunto, $\{x_1, x_2, \dots, x_n\}$, por n vezes com reposição.

1 Veja-se Efron(1979) pp. 5 e 6.

- 2) Para esta realização calcula-se o valor da estatística,

$$\hat{\theta}' = \hat{\theta}(x'_1, x'_2, \dots, x'_n) ;$$

- 3) Repetem-se os procedimentos 1) e 2) um número suficientemente grande, B, de vezes.

O conjunto de B valores, $\hat{\theta}'_1, \dots, \hat{\theta}'_b, \dots, \hat{\theta}'_B$, assumidos pela v.a., $\hat{\theta}'(X'_1, X'_2, \dots, X'_n)$, obtidos através deste procedimento, permite calcular as estimativas pretendidas.

A estimativa *bootstrap* de $E_{\theta}(\hat{\theta})$ será igual a,

$$E_{\theta}(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^B \hat{\theta}'_b = \hat{\theta}'(.) . \quad (2.2)$$

A partir de (2.2) é possível obter a estimativa do enviesamento de $\hat{\theta}$, $\xi'(\hat{\theta})$, de acordo com,

$$\xi'(\hat{\theta}) = \hat{\theta}'(.) - \hat{\theta}(x_1, x_2, \dots, x_n) . \quad (2.3)$$

A estimativa do desvio padrão de $\hat{\theta}$ será igual a

$$\hat{\sigma}'(\hat{\theta}) = \left\{ \frac{1}{B-1} \sum_{b=1}^B [\hat{\theta}'_b - \hat{\theta}'(.)]^2 \right\}^{1/2} . \quad (2.4)$$

A distribuição estimada *bootstrap* de $\hat{\theta}$, $\hat{G}(s)$, coincide com as frequências relativas acumuladas dos B valores *bootstrap* $\hat{\theta}'_1, \dots, \hat{\theta}'_b, \dots, \hat{\theta}'_B$, obtidos com o procedimento Monte Carlo apresentado, donde,

$$\hat{G}(s) = \frac{1}{B} \# \{ \hat{\theta}'_b < s \} . \quad (2.5)$$

Assim a estimativa de $P_F\{\hat{\theta} < k\}$ é igual a,

$$P_F\{\hat{\theta}^* < k\} = \frac{\# \{\hat{\theta}_b^* < k\}}{B}.$$

Quando $B \rightarrow \infty$ as aproximações dadas pelo algoritmo Monte Carlo tendem para as estimativas *bootstrap* teóricas. A escolha do valor de B tem de obedecer fundamentalmente a dois critérios: por um lado é desejável que B seja suficientemente grande para que as estimativas *bootstrap* numéricas se ajustem convenientemente às estimativas teóricas ; mas, por outro lado, há que limitar o valor de B de forma a não tornar os custos calculatórios incomportáveis.

Efron(1987) apresenta um critério para determinar o valor de B baseado no coeficiente de variação da estimativa, CV, com CV igual à razão entre o desvio padrão e o valor esperado. Distingue o valor de B conforme se pretenda estimar o desvio padrão e o enviesamento da estatística ou a sua distribuição amostral, sendo neste último caso o valor de B bastante superior.

O coeficiente de variação, não condicional (traduzindo assim a média sobre todas as possíveis realizações \underline{x}), da estimativa do desvio padrão é aproximadamente igual a

$$CV(\hat{\sigma}^*) = \left\{ CV(\hat{\sigma})^2 + \frac{E\{\hat{\delta}\} + 2}{4B} \right\}^{1/2},$$

com $\hat{\delta}$ o kurtosis da distribuição *bootstrap* de $\hat{\theta}^*$ dado \underline{x} . O autor conclui em estudos de simulação que a partir de B superior a 100 e com $CV(\hat{\sigma}) > 0.1$, como sucede em geral, não se verificam diferenças sensíveis entre $CV(\hat{\sigma})$ e $CV(\hat{\sigma}^*)$. Como tal, na determinação de estimativas *bootstrap* do desvio padrão ou enviesamento é geralmente suficiente um valor de B=100.

A fixação do valor de B no cálculo da estimativa da distribuição de $\hat{\theta}$ passa pela análise do coeficiente de variação condicional de $\hat{\theta}^*(\alpha) - \hat{\theta}$, com $\hat{\theta}^*(\alpha)$ o $\alpha \times 100$ -ésimo percentil de $\hat{C}(s)$. Este coeficiente, supondo que a distribuição *bootstrap* teórica é aproximadamente normal, verifica

$$CV\{\hat{\theta}^*(\alpha) - \hat{\theta} \mid \underline{x}\} \approx \frac{1}{\sqrt{B} |z_\alpha|} \left\{ \frac{\alpha(1-\alpha)}{\phi(z_\alpha)^2} \right\}^{1/2},$$

sendo $\phi(z)$ a densidade da $\mathcal{N}(0,1)$. Efron conclui através de alguns estudos de simulação que o valor aconselhado para B é no mínimo $B=1000$.

2.2 - REAMOSTRAGENS ALTERNATIVAS

O *bootstrap* é um processo de reamostragem na medida em que na sua aplicação é necessário gerar artificialmente novas amostras a partir da amostra inicial. No *bootstrap* convencional, atrás descrito, as reamostragens são definidas pressupondo que as variáveis *bootstrap* tem distribuição F , a distribuição empírica da amostra. Contudo não tem de ser sempre esta, necessariamente, a estrutura de reamostragem.

Como F é uma distribuição discreta, as amostras construídas a partir de F têm propriedades muito específicas. Todos os seus valores pertencem ao conjunto de valores da amostra inicial, como tal, todas as reamostras de composição diferente da amostra observada terão de conter valores repetidos, podendo originar amostras bastante desequilibradas.

Quanto menor for n , a dimensão da amostra inicial, mais reduzido será o número de combinações diferentes das n observações e dado que B é necessariamente elevado, haverá nas B amostras

bootstrap uma maior probabilidade de obter amostras repetidas no sentido que geram o mesmo valor para a estatística, conduzindo a um suporte da distribuição *bootstrap* da estatística bastante discreto.

Os processos alternativos de reamostragem - *bootstrap* equilibrado, *bootstrap* alisado, e o *bootstrap* paramétrico - superam estas especificidades, o primeiro impedindo a predominância da repetição desequilibrada de algumas observações, e os restantes permitindo que as amostras artificiais sejam muito mais variadas, incluindo valores para além dos observados. No fundo, o *bootstrap* alisado e o *bootstrap* paramétrico permitem estimar a distribuição do universo F por uma função mais suave do que \hat{F} . seja \tilde{F} . O efeito líquido deste processo resume-se a estimar a funcional $\theta(F)$ por $\theta(\tilde{F})$, enquanto que no *bootstrap* convencional $\theta(F)$ é estimada por $\theta(\hat{F})$, tal como na teoria da M.V. não paramétrica.¹

Seja \tilde{F} , a função alternativa de reamostragem. O *bootstrap*, com esta função de reamostragem processa-se da mesma forma que o *bootstrap* convencional, tendo apenas em consideração no algoritmo *bootstrap*, ponto 1), que as reamostras são geradas a partir da estrutura de probabilidade caracterizada por \tilde{F} . isto é, cada amostra é constituída por n concretizações independentes da v.a. com distribuição \tilde{F} .

¹ A vantagem do *bootstrap* em relação à teoria da M.V. não paramétrica reside na sua possibilidade de apresentar soluções através de aproximações de Monte Carlo para os casos onde a dedução de $\theta(\hat{F})$ ou a inferência não numérica sobre $\theta(\hat{F})$ é impossível ou demasiado complexa.

2.2.1 - *Bootstrap* equilibrado

As estimativas numéricas obtidas com o método de Monte Carlo apresentam erros de simulação intimamente relacionados com a estrutura desequilibrada da amostra e que podem ser importantes quando B não é muito elevado.

Para reduzir o erro de simulação na obtenção de estimativas *bootstrap* do enviesamento, da variância e da distribuição de $\hat{\theta}$, tornando a simulação mais eficiente, Davidson, Hinkley e Schechtman(1986) preconizam uma estrutura de reamostragem de forma a produzir uma simulação *bootstrap* equilibrada.¹ A simulação diz-se equilibrada na medida em que se exige que no total das B reamostras cada observação x_i apareça um número igual de vezes.

O método processa-se juntando B cópias da amostra observada \underline{x} numa fila com o total de nB elementos. De seguida, faz-se uma permutação aleatória aos seus elementos. Depois retiram-se n elementos ordenadamente de cada vez por B vezes, constituindo as B amostras *bootstrap*. Formalizando o processo, refere-se que os B vectores de reamostragem $\underline{n}P_b^* = n[P_{b1}^*, \dots, P_{bn}^*]$ formam uma matriz com distribuição hipergeométrica em que a soma ao longo de todas as linhas é igual a n e a soma ao longo de todas as colunas é igual a B .

Os autores mostram que para aproximações lineares e quadráticas da estatística, este método reduz a média e a variância do erro de simulação das estimativas do enviesamento e variância de $\hat{\theta}$.

¹ Os autores apontam ainda um método alternativo a aplicar quando a estatística verifica uma aproximação linear com momentos conhecidos, sendo apenas necessário simular o residual desta aproximação. Esta via é particularmente útil quando se utiliza a aproximação a uma v.a. normal ou em aproximações de ordem superior como a aproximação em Expansão de Edgeworth.

2.2.2 - *Bootstrap* alisado

Efron(1979) introduz o *bootstrap* alisado considerando que a função de distribuição que estrutura a reamostragem, \hat{F} , consiste num estimador de F alisado, em janela, com a média e a variância de \hat{F} . Em estudos posteriores, Efron fixa para \hat{F} a distribuição resultante da convolução de \hat{F} com uma normal ponderada por um factor de escala, onde a média e a variância da normal são iguais aos respectivos momentos da amostra. Esta via é indicada quando não se pretendem abandonar completamente as hipóteses tradicionais da normalidade da população.

Outros autores introduziram formas mais sofisticadas para suavizar \hat{F} , das quais se destacam duas.¹ Estas duas formas consistem em considerar para \hat{F} respectivamente uma estimativa *kernel* e uma estimativa *kernel shrunk* de F . Para determinar estas estimativas escolhe-se uma função *kernel*, K , de forma que K seja uma função de densidade de probabilidade simétrica com variância unitária. K pode ser por exemplo a Normal estandardizada. A estimativa *kernel* da densidade da população do universo, f , para uma população k -dimensional, é igual a

$$\hat{f}_h(x) = |V|^{-1/2} n^{-1} h^{-k} \sum_{i=1}^n K\{h^{-1} V^{-1/2}(x - X_i)\} \quad (2.6)$$

onde V representa a matriz de covariâncias de X_i (k -dimensional). Esta matriz pode ser substituída por um estimador consistente

¹ Hall e outros(1989) pp. 694 e 695 apresentam ainda formas alternativas de suavização de \hat{F} .

como a matriz das covariâncias amostrais.¹ Por sua vez, a estimativa *kernel shrunk* de f , para uma população k -dimensional, é dada por

$$\hat{f}_{h,s}(x) = (1+h^2)^{k/2} \hat{f}_h \{(1+h^2)^{k/2} x\} . \quad (2.7)$$

Por fim, $\hat{F}(x)$ vem igual a $\int_{-\infty}^x \hat{f}_h(y)dy$ ou $\int_{-\infty}^x \hat{f}_{h,s}(y)dy$ conforme se utilize (2.6) ou (2.7).

Para problemas em que $\theta(F)$ é uma funcional linear em F , ou localmente linearizável, Silverman e Young(1987) mostram que o *bootstrap* alisado através de (2.6) ou (2.7) é mais vantajoso do que o *bootstrap* convencional na medida que permite que $\theta(\hat{F})$ tenha um erro quadrático médio inferior ao erro quadrático médio de $\theta(F)$, desde que se escolha um valor positivo conveniente para h . No entanto, os autores não especificam quais os valores ótimos para h , referindo apenas que a *performance* do *bootstrap* alisado está bastante dependente de "quanto se alisa", constituindo este assunto um problema complexo. Para ultrapassar a questão é sempre possível ensaiar vários valores de h e escolher aquele que conduza a melhores resultados.

Hall, Diccicio e Romano(1989) apontam que a diminuição do erro quadrático médio permitida pelo *bootstrap* alisado não é válida assintoticamente quandoquando o estimador é \sqrt{n} -consistente, visto que, nesta situação, $\theta(\hat{F})$ e $\theta(\hat{F})$ têm assintoticamente igual variância. É o caso das estatísticas que podem ser expressas como funções diferenciáveis de médias de vectores como médias, razões entre médias, variâncias, razões entre variâncias, coeficientes de correlação, entre outras.

Os autores mencionados concluem que:²

1 Não é forçoso impôr variância unitária a K e a \hat{f}_h .

2 Hall, Diccicio e Romano(1989) p.694

When the quantity of interest is genuinely nonparametric in character, and cannot be estimated \sqrt{n} -consistently, smoothing the bootstrap can substantially improve convergence rates;

conclusão que testam na estimação de quantis.

Com o *bootstrap* convencional a estimativa da variância do quantil da amostra sofre de um erro de ordem $n^{-1/4}$ quando a dimensão da amostra, n , aumenta. Os autores provam que utilizando um estimador *kernel* de ordem $r \geq 2$, com $\int |x^r K(x)| dx < \infty$, escolhendo h adequadamente a ordem do erro pode ser manipulada para $n^{-r/(2r+1)}$, marcadamente melhor que $n^{-1/4}$ qualquer que seja o valor de r . Apresentam ainda uma fórmula para a suavização óptima.

Young(1988) realiza um estudo sobre a performance do *bootstrap* alisado na estimação do desvio padrão do coeficiente de correlação. Conclui que a suavização é proveitosa ao diminuir o erro quadrático médio do estimador, particularmente quando a população é normal (influindo aqui o facto de ter utilizado para *kernel* a Normal estandardizada). Os efeitos benéficos diminuem com o aumento da dimensão da amostra (o que está de acordo com a igualdade assintótica verificada para este tipo de estatísticas entre os momentos obtidos com o *bootstrap* convencional e os obtidos com o *bootstrap* alisado, conforme foi já referido). No entanto apesar de diminuir o erro quadrático médio influi no agravamento do enviesamento do estimador. O autor afirma que o grau de suavização depende da distribuição da população e determina empiricamente o grau óptimo através da minimização numérica da função de erro quadrático médio estimado.

Do exposto é de realçar que o *bootstrap* alisado mostra-se vantajoso em relação ao *bootstrap* convencional, em muitas situações, por diminuir o erro quadrático médio do estimador em pequenas amostras. Em certos casos, possibilita ainda uma

aceleração da taxa de convergência do estimador da variância da estatística. Na suavização é importante a escolha do *kernel* e a determinação do h ótimo, dependendo ambos da distribuição subjacente à população. Seria interessante desenvolver pesquisas no sentido de determinar funções *kernel* robustas. O ensaio de várias hipóteses para K , quando possível, e a determinação empírica de h apresentada por Young conduzirão a resultados mais seguros.

2.2.3 - *Bootstrap* paramétrico

O *bootstrap* paramétrico deve ser implementado quando existe um conhecimento seguro sobre a distribuição do universo (ignorando-se obviamente os valores dos parâmetros). A reamostragem será processada a partir da distribuição postulada para a população sendo os valores dos parâmetros substituídos por estimativas consistentes dos verdadeiros parâmetros.

Existindo uma boa aderência entre a distribuição fixada para as variáveis *bootstrap* e a família de distribuições do universo, o *bootstrap* paramétrico é conceptualmente mais correcto que o convencional ou o alisado. Caso contrário, os resultados podem ser desastrosos, sendo preferível utilizar o *bootstrap* alisado ou mesmo o convencional que de certa forma apresentam uma maior robustez.

2.3 - PROPRIEDADES ASSIMPTOTICAS DO *BOOTSTRAP*

2.3.1 - Convergência da média *bootstrap*.

Seja $\bar{X}_n^* = (1/n) \sum_{i=1}^n X_i^*$ e $\hat{s}_n^* = (1/n) \sum_{i=1}^n (X_i^* - \bar{X}_n^*)^2$. Bickel e Freed-

man(1981) provam que, quando as variáveis possuem variância finita positiva, para quase todas as realizações da amostra,

i) A distribuição condicionada¹ de $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)$ converge (em lei) para a $\mathcal{N}(0, \sigma^2)$, sendo $\sigma^2 = \text{Var}(X_1)$;

ii) \hat{s}_n^* tende em probabilidade condicionada para σ .

Estes resultados verificam-se ainda para v.a. k-dimensionais. Se a reamostragem for definida pressupondo uma função de distribuição diferente da empírica, seja F_n , i) e ii) são válidos desde que F_n convirja em distribuição para F e $\int x^2 dF_n(x)$ tenda para $\int x^2 dF(x)$.

Como $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)$ tem assintoticamente a mesma distribuição que $\sqrt{n}(\bar{X}_n - \mu)$, a $\mathcal{N}(0, \sigma^2)$, é possível concluir que, em grandes amostras, a estimativa *bootstrap* da média aproxima de forma correcta a média da população.

¹ Daqui por diante condicionada refere-se a condicionada pela amostra.

Singh(1981) reforça esta conclusão provando que, sendo $E\{X_1^2} < \infty$, então,¹

$$\sup_{x \in \mathbb{R}} \left| P\{\sqrt{n}(\bar{X}_n - \mu) \leq x\} - P\{\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x\} \right| \rightarrow 0 \text{ q.c. .}$$

Ainda segundo Singh(1981) é possível concluir que verificando-se $E\{X_1^3} < \infty$ a taxa de convergência desta aproximação é $O_p(n^{-1/2}(\log \log n)^{1/2})$.

Este autor mostra ainda que para $E\{|X|^3} < \infty$ e F não recticulada a convergência entre as distribuições de $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ e de $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)/\hat{s}_n$ processa-se em termos de $O_p(n^{-1/2})$. Esta convergência é mais rápida do que a convergência entre $\sqrt{n}(\bar{X}_n - \mu)$ e $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)$, que como se viu era $O_p(n^{-1/2}(\log n \log n)^{1/2})$. Como tal, a *standardização* prévia das estatísticas *bootstrap* pode potenciar a performance dos resultados *bootstrap* dado que assim se obtém uma distribuição mais próxima da verdadeira distribuição da estatística.

2.3.2 - Convergência de estatísticas *bootstrap studentizadas*

A estatística *studentizada* mais trivial é a conhecida $\sqrt{n}(\bar{X}_n - \mu)/\hat{s}_n$. De acordo com o teorema clássico do Limite Central, esta estatística tem distribuição assintótica $\mathcal{N}(0,1)$. Atendendo aos resultados i) e ii) acima indicados é possível concluir que $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)/\hat{s}_n^*$ tende igualmente (quando $n \rightarrow \infty$) para a $\mathcal{N}(0,1)$. Dado que a estatística *bootstrap* e a estatística clássica tem igual distribuição limite, a $\mathcal{N}(0,1)$, fica patente a validade assintótica do *bootstrap* de estatísticas *studentizadas*.

¹ Neste capítulo q.c. significa "para quase todas as realizações da amostra".

Em Babu e Singh(1983) analisa-se a convergência de uma estatística *studentizada* mais geral, pressupondo um problema onde coexistem k amostras provenientes de k populações diferentes, em que a população i se caracteriza pela distribuição F_i . A estatística considerada é igual a $t_N = (\hat{\theta}_N - \theta) / \hat{s}_N$, e θ é uma combinação linear de médias,

$$\theta = \sum_{i=1}^k l_i \mu_i, \quad l_i \neq 0 \quad \text{e} \quad \hat{\theta}_N = \sum_{i=1}^k l_i \bar{X}_i,$$

em que μ_i denota o centro de gravidade de F_i , e \bar{X}_i a média da amostra de dimensão n_i ($\sum_{i=1}^k n_i = N$), retirada da população F_i .

e $\hat{s}_N^2 = 1/N \sum_{i=1}^k l_i^2 \hat{s}_i^2$. A congénere *bootstrap* de t_N , t_N^* , é definida por

$t_N^* = (\hat{\theta}_N^* - \hat{\theta}_N) / \hat{s}_N^*$, onde $\hat{\theta}_N^*$ e \hat{s}_N^* são obtidos da mesma forma que $\hat{\theta}_N$ e \hat{s}_N substituindo os X_i por X_i^* .

Aqueles autores mostram que sendo finito o sexto momento de F_i , $1 \leq i \leq k$ e F_i contínua, admitindo-se ainda que os n_i são de dimensão comparável, a convergência entre as distribuições de t_N^* e t_N processa-se à taxa $o_p(N^{-1/2})$, desde que $B/(N \log N) \rightarrow \infty$. Afirmam que esta convergência é mais rápida que a convergência entre as distribuições de $(\hat{\theta}_N - \theta)$ e $(\hat{\theta}_N^* - \hat{\theta}_N)$. Verificam também que, nas condições anteriores, o quantil de ordem t da distribuição *bootstrap* converge assintoticamente para o quantil de igual ordem da distribuição de t_N , sendo a diferença entre ambos $o_p(N^{-1/2})$.

Babu e Singh(1983) estabelecem ainda a validade da aproximação à distribuição *bootstrap* obtida com as técnicas Monte Carlo. Seja $H_{N,B}^*$ a aproximação à distribuição *bootstrap* teórica de t_N^* , H_N^* , determinada com o algoritmo baseado nas técnicas de Monte Carlo. Desde que $B/N \log N \rightarrow \infty$ com $N \rightarrow \infty$ e sendo finito o sexto momento de F_i , $1 \leq i \leq k$ com F_i contínua, então verifica-se,

$$\sqrt{N} \sup_{x \in \mathbb{R}} |\hat{H}_{N,B}^*(x) - \hat{H}_N^*(x)| \rightarrow 0 \text{ q.c.},$$

$$\sqrt{N} |\hat{H}_{N,\beta}^{*-1}(t) - \hat{H}_N^{*-1}(t)| \rightarrow 0 \text{ q.c.},$$

onde se destaca que o erro de ambas as aproximações é da ordem $o_p(N^{-1/2})$.

2.3.3 - Convergência do processo empírico *bootstrap*

Sabe-se que $\sqrt{n}(\hat{F}_n - F)$ converge em distribuição para $B(F)$, com $B(F)$ uma ponte *Browniana* em $[0,1]$, assegurando que, assintoticamente \hat{F}_n aproxima a distribuição do universo F .

Bickel e Freedman(1981) demonstram o análogo *bootstrap* desta assumpção, isto é, mostram que $\sqrt{n}(\hat{F}_n^* - \hat{F}_n)$ com \hat{F}_n^* a distribuição empírica da amostra *bootstrap*, tende em distribuição para $B(F)$ (com $B(F)$ conforme foi indicado).

Conjugando estes resultados, conclui-se que a distribuição empírica da amostra *bootstrap*, para n suficientemente grande, aproxima a distribuição do universo, F .

Para uma estrutura de reamostragem alternativa a partir de uma \hat{F}_n geral, mantém-se a validade do resultado exposto desde que \hat{F}_n tenda quase certamente para F .

2.3.4 - Convergência do processo quantil *bootstrap*

O processo quantil é definido em $[0,1]$ por $Q_n(t) = \sqrt{n} \{ \hat{F}_n^{-1}(t) - F^{-1}(t) \}$. Sabe-se que assintoticamente $\hat{F}_n^{-1}(t)$ aproxima em probabilidade $F^{-1}(t)$.

A réplica *bootstrap* de $Q_n(t)$ define-se por,

$$Q_n^*(t) = \sqrt{n} \{ \hat{F}_n^{*-1}(t) - \hat{F}_n^{-1}(t) \}.$$

Bickel e Freedman(1981) demonstram que assintoticamente $Q'_n(t)$ converge para a distribuição limite de $Q_n(t)$. Este resultado revela que em grandes amostras o quantil de uma dada ordem da distribuição empírica *bootstrap* aproxima o quantil (de igual ordem) de F .

Singh(1981) mostra também a consistência da aproximação *bootstrap* à distribuição de $Q_n(t)$, deduzindo a taxa exacta à qual a diferença entre as duas distribuições converge para zero. Assim, a partir do teorema 2 do referido texto, que exige como condição que F tenha uma derivada de segunda ordem limitada na vizinhança de $F^{-1}(t)$ e que seja $F'[F^{-1}(t)] > 0$, é possível concluir que esta taxa é $O_p(n^{-1/4}(\log \log n)^{1/2})$.

O autor extrai deste teorema uma ilacção de grande interesse - quando $F'(F^{-1}(t))$ (com F' a primeira derivada de F) é substituído por uma estimativa $O_p(n^{-1/4}(\log \log n)^{1/2})$ baseada na amostra, a aproximação *bootstrap* de $Q_n(t)$ não é inferior à aproximação dada pela Normal.

Com base na validade do *bootstrap* para o processo quantil determinam-se algumas aplicações interessantes, das quais se menciona a construção de I.C. para a mediana ou para pivôs definidos a partir das médias "aparadas" ou de variâncias *Win-sorizadas*. No entanto, para o caso da mediana, Ghosh et al(1984) mostram que por vezes o *bootstrap* pode não resultar por a estimativa *bootstrap* do d.p. da estatística, em alguns casos, tender para infinito. Para que se verifique no caso da mediana o teorema demonstrado por Bickel e Freedman é necessário impôr algumas condições aos momentos de F , nomeadamente $E|X_1|^\alpha < \infty$ para um dado $\alpha > 0$. Os autores propõem uma robustificação do *bootstrap* de forma a não ser exigida esta condição para verificar a consistência assintótica do desvio padrão *bootstrap* da mediana. A

robustificação consiste em *Winsorizar* ou em alternativa "apapar" a amostra inicial. Para mais detalhes sobre estes processos veja-se Ghosh et al(1984) p.1134

Em Bickel e Freedman(1981) mostra-se ainda a validade do *bootstrap* para combinações lineares de estatísticas de ordem.

2.3.5 - Superioridade do *bootstrap* sobre a aproximação à Normal

A superioridade da aproximação *bootstrap* à distribuição de uma estatística relativamente à aproximação à normal será analisada recorrendo em grande parte dos casos à expansão de Edgeworth da função de distribuição, $F(x)$, de uma v.a..

Deste modo, (supondo para simplificação da formalização que $F(x)$ está estandardizada ou *studendizada*), $F(x)$ é desenvolvida numa série em que o primeiro termo é uma Normal estandardizada e os outros termos envolvem sucessivas potências de $1/n$ de acordo com,

$$F(x) = \Phi(x) + \int_{-\infty}^x \sum_{r=1}^{s-2} n^{-r/2} p_r(x) \phi(x) dx \quad (2.10)$$

Os coeficientes de p_r são polinómios nos momentos da variável aleatória de ordem inferior ou igual a s .¹

Quando uma dada distribuição verifica a expansão (2.10) então a aproximação dada pela Normal sofre um erro de grandeza $O(n^{-1/2})$, pois a partir de (2.10) $F(x) = \Phi(x) + O(n^{-1/2})$. Uma aproximação superior será aquela que conseguir captar correctamente termos

¹ Para que exista a E.E. de ordem s é necessário que $F(x)$ seja fortemente não quadriculada, existam e sejam finitos os momentos de $F(x)$ até a ordem s , sendo também necessário a verificação da condição de Cramér,

$$\lim_{|t| \rightarrow \infty} \left| \int e^{itx} dF(x) \right| < 1$$

de ordem superior da E.E. referida. De uma forma geral, a aproximação que consiga captar inteiramente o termo que envolve $n^{-k/2}$ sofre de um erro $o(n^{-k/2})$, ou $O(n^{-(k+1)/2})$.

Beran(1982) analisa a superioridade do *bootstrap* utilizando a teoria de decisão estatística. O autor considera a estatística \hat{R}_n , estimador da funcional $R(F)$, definida de forma que a distribuição assintótica de $\sqrt{n}(\hat{R}_n - R(F))$ seja uma Normal com média zero.

Sob algumas condições de regularidade o autor prova que a distribuição *bootstrap* de $\sqrt{n}(\hat{R}_n - R(F))$, \hat{H}_n^* , sendo a distribuição da estatística *bootstrap* $\sqrt{n}(\hat{R}_n^* - \hat{R}_n(F))$, é assintoticamente minimax, onde a função perca é uma qualquer função monótona crescente limitada de uma dada norma da diferença $\sqrt{n}\{H_n^*(x) - H_n(x)\}$ ponderada, com $H_n(x)$ a verdadeira distribuição de $\sqrt{n}(\hat{R}_n - R(F))$. Este resultado é ainda válido quando a reamostragem é feita a partir de uma distribuição mais geral F_n desde que,

$$\limsup_{n \rightarrow \infty} P(\sqrt{n} \|\hat{F}_n - \hat{F}_n\| > \epsilon) = 0,$$

para qualquer ϵ , qualquer c positivo, e qualquer F em \mathcal{F} (\mathcal{F} é o conjunto de todas as distribuições com suporte num dado intervalo compacto fixo).

A aproximação normal $\Phi(x/\hat{s}_n)$, (pressupondo σ desconhecido) não é geralmente assintoticamente minimax porque é enviesada dado que a estimativa de $H_n(x)$ assintoticamente minimax é a Normal $\Phi(x/\hat{s}_n)$ subtraída da estimativa natural do enviesamento de $\sqrt{n}(\hat{R}_n - R(F))$.

Liu e Singh(1987) exploram com mais pormenor este problema. Mostram que, para todos os casos em que existe E.E. de $H_n(x)$, se $\sqrt{n}(\hat{s}_n^2 - \sigma^2)$ tem como distribuição limite uma $\mathcal{N}(0, a^2)$ (com a uma dada constante) então $\sqrt{n}\{H_n(x) - \hat{H}_n^*(x)\}$ tem como distribuição assintótica uma $\mathcal{N}(0, b^2)$, com $b^2 = \phi^2(x/\sigma)x^2\alpha^2/\sigma^6$, enquanto que a

distribuição assintótica de $\sqrt{n}\{H_n(x) - \Phi(x/\xi_n)\}$ é uma $\mathcal{N}(q_1(x)\phi(x/\sigma), b^2)$.¹ Assim, a aproximação dada pela Normal sofre de um enviesamento igual a $q_1(x)\phi(x/\sigma)$. Desta forma o *bootstrap* é superior à Normal em termos de erro quadrático médio assintótico. Esta superioridade - que os autores denominam por correcção parcial do termo $n^{-1/2}$ - depende geralmente do enviesamento de \hat{R}_n como estimador de $R(F)$ e da assimetria de $H_n(x)$.

Há ainda a notar que a correcção parcial dada pelo *bootstrap* melhora a estimativa dos quantis de $H_n(x)$ no sentido de que reduz o erro quadrático médio assintótico.

É de realçar que para estatísticas *studentizadas* verifica-se um fenómeno de correcção total, pois, como foi apontado, a diferença entre a verdadeira distribuição e a distribuição *bootstrap* desta estatística é $o_p(n^{-1/2})$ o que traduz um ajustamento totalmente correcto do termo em $n^{-1/2}$ da E.E. de ambas as distribuições.

Por outro lado, Beran(1982) mostra que a estimativa dada pela E.E. empírica de segunda ordem (obtida substituindo os momentos teóricos do termo em $n^{-1/2}$ pelos momentos empíricos) é também assintoticamente minimax, e é equivalente à estimativa *bootstrap* $H_n^*(x)$ até aos termos de ordem $n^{-1/2}$.

Mais recentemente, Bhattacharya e Qumsiyeh(1989) mostram que o *bootstrap* pode ainda ser superior à E.E. empírica de dois termos. Desenvolvem a sua análise para estatísticas *studentizadas*, escalares de \mathbb{R} , do tipo $T_n = \sqrt{n}H(\bar{X})$ em que H é uma função com derivadas até à ordem s num conjunto aberto que contém $E(X_1)$ e

¹ Onde $q_1(x) = \int p(x)dx$, (com $p_1(x)$ o polinómio do termo em $n^{-1/2}$ da E.E. dada por (2.10) aplicada a $H_n(x)$).

tal que $H[E(X_1)] = 0$, condições que garantem que a distribuição de T_n verifique uma E.E. de ordem $(s-1)$. Seja,

$$q_2(x)\phi(x) = \int_{-\infty}^x p_2(x)\phi(x) dx,$$

com $p_2(x)$ o polinómio do terceiro termo da E.E. da densidade *bootstrap* de T_n .

Aqueles autores mostram que, verificando-se determinadas condições de regularidade, para todas as funções-L de perca, limitadas, simétricas e *bowl-shaped*, sendo $q_2(x) \neq 0$ verifica-se a seguinte desigualdade para todo o x ,¹

$$\lim_{n \rightarrow \infty} EL\{n[P.(T_n^* \leq x) - P(T_n \leq x)]\} < \lim_{n \rightarrow \infty} EL\left\{n\left[\phi(x) + \int_{-\infty}^x p_{1,n}(x)\phi(x) dx - P(T_n \leq x)\right]\right\}.$$

Quando $q_2(x) = 0$ a desigualdade transforma-se numa igualdade.

Estas conclusões são ainda válidas quando se avaliam as aproximações com as normas L_p ,² registando-se,

$$\lim_{n \rightarrow \infty} n\|P.(T_n^* \leq x) - P(T_n \leq x)\|_p < \lim_{n \rightarrow \infty} n\left\|\phi(x) + \int_{-\infty}^x p_{1,n}(x)\phi(x) dx - P(T_n \leq x)\right\|_p.$$

Para $q_2(x) = 0$ verifica-se a igualdade das duas expressões.

Estes resultados realçam o potencial do *bootstrap* dado que assintoticamente a sua distribuição está mais próxima da distribuição da estatística do que a distribuição resultante da E.E. empírica de dois termos que por sua vez já constitui uma aproximação superior á normal.

¹ L em \mathbb{R}^p é *bowl-shaped* se $\{x \in \mathbb{R}^p : L(x) \leq c\}$ é convexo para qualquer c ; é simétrica se $L(x) = L(-x)$ para qualquer x ; e constitui uma função de perca se $L(0) = 0$, $L(x) > 0$ para qualquer $x \neq 0$.

² A norma L_p é igual a, $\|Y\|_p = (E\{|Y|^p\})^{1/p}$.

Até agora, as taxas apresentadas de convergência da aproximação *bootstrap* verificam-se apenas em probabilidade quase certamente, ao contrário da Normal que é $O(n^{-1/2})$. Babu e Bose [(1989a) e (1989b)] vêm enriquecer o conhecimento das características da aproximação *bootstrap* deduzindo taxas de convergência que são válidas com uma probabilidade que se aproxima rapidamente de um, podendo ser assim mais directamente comparáveis com a taxa da normal.

Os autores centram o seu estudo em estatísticas *studentizadas* com a seguinte forma,

$$T_n = \sqrt{n} \frac{g(\bar{X}) - g(\mu)}{v \left[(1/n) \sum_{i=1}^n r(X_i) \right]}, \quad (2.11)$$

onde $X_i \in \mathbb{R}^k$, g é uma função de \mathbb{R}^k em \mathbb{R} , r é uma função de \mathbb{R}^k em \mathbb{R}^i e v de \mathbb{R}^i em \mathbb{R} . Sabe-se ainda que $v(E_f r(X_i)) = \sigma$ e $v(E_{f_n} r(X_i)) = \hat{\sigma}_n$. A estatística t de Student é um caso particular de (2.11).

Babu e Bose têm como objectivo deduzir a grandeza do erro entre α e a verdadeira probabilidade (a probabilidade na verdadeira distribuição de T_n) do quantil α da distribuição *bootstrap* de T_n , a distribuição amostral de T_n^* , H_n^* , com T_n^* a versão *bootstrap* de T_n .

Analise-se o caso particular em que σ é conhecido, assumindo T_n a forma $\sqrt{n}(g(\bar{X}) - g(\mu))/\sigma$ e T_n^* será $\sqrt{n}(g(\bar{X}^*) - g(\bar{X}))/\hat{\sigma}_n$. Supondo que $g(x)$ é linear em x , Babu e Bose concluem que, perante algumas condições de regularidade (que garantem a existência de E.E. de ordem 12), se verifica

$$\sup_{0 \leq \alpha \leq 1} |P(T_n < H_n^{*-1}(\alpha)) - \alpha| = O(n^{-1}(\log n)^{1/2}) \quad (2.12)$$

Quando $g(x)$ é não linear este resultado é ainda válido desde que g seja diferenciável até à ordem 4.

Para σ desconhecido, o resultado (2.12) continua válido desde que g e v sejam diferenciáveis respectivamente até à quarta ordem e até à terceira ordem.

A constatação (2.12) traduz uma superioridade do *bootstrap* em relação à aproximação à normal na medida em que a grandeza do erro do quantil estimado através da distribuição *bootstrap* da estatística *studentizada* é menor do que a estimativa dada pelo quantil de igual ordem da normal estandardizada.

Babu e Bose(1989b) analisam ainda a grandeza do erro da aproximação *bootstrap* à verdadeira distribuição de T_n . Assim sob determinadas condições de regularidade (que ressaltam a existência de E.E. de ordem $3+\gamma, \gamma>3$), os autores mostram que

$$P(\sup_x |P_F(T_n \leq x) - P_*(T_n^* \leq x)| \geq n^{-1/2} \epsilon_n) = o(n^{-(1+\delta)}) \quad (2.13)$$

para algum $\epsilon_n \rightarrow 0$ e $\delta > 0$. Neste resultado é patente que a grandeza do erro entre a distribuição *bootstrap* e a verdadeira distribuição é menor do que a existente entre esta distribuição e a respectiva distribuição limite Normal.

Este problema pode ser abordado numa perspectiva mais geral. Seja $\hat{R}_n = \hat{R}_n(X_1, \dots, X_n)$ uma dada estatística em que $X_i \in \mathbb{R}^k$, e H_n representa a verdadeira distribuição de $\hat{R}_n - R_n(F)$ (com $\{R_n(F)\}$ uma sequência de constantes), sendo H_n^* a aproximação *bootstrap* de H_n . Babu e Bose(1989b) estabelecem que sob determinadas condições sobre \hat{R}_n e sobre os momentos de F (com F suave), sendo X_i ($i=1, \dots, n$) variáveis fortemente não recticuladas verifica-se,

$$P(\sup_x |H_n(x) - H_n^*(x)| \geq \epsilon_n n^{-1/2}) = O(n^{-(1+\delta)}) \quad , \quad (2.14)$$

com $\epsilon \rightarrow 0$ uma dada sequência e $\delta > 0$ depende das hipóteses fixadas

sobre os momentos. A partir de (2.14) é possível concluir que o *bootstrap* apresenta maiores potencialidades na aproximação da distribuição de uma dada estatística do que a aproximação clássica à Normal.

Para extrair os pontos de confiança de intervalos de confiança para o parâmetro θ , considerando a técnica *bootstrap*:

A dedução de um intervalo de confiança para θ pode ser feita através de um procedimento análogo ao utilizado para a extração de pontos de confiança para o parâmetro θ , considerando a técnica *bootstrap*.

$$H(\alpha) = P(\hat{\theta} - \hat{\theta}^* \leq \alpha) \quad (3.1)$$

$$K(\alpha) = P(\hat{\theta} - \hat{\theta}^* \leq \alpha) \quad (3.2)$$

onde $\hat{\theta}$ é a estimativa de θ de n observações e $\hat{\theta}^*$ é a estimativa de θ de n^* observações. A expressão (3.1) representa uma aproximação de θ por meio da informação de n observações.

De acordo com a expressão (3.1) para $\alpha = 0$, são conhecidos os pontos de confiança de θ que verificam a propriedade: quando o tamanho da amostra n tende a infinito, o ponto crítico α tende a zero.

$$H(\alpha) = P(\hat{\theta} - \hat{\theta}^* \leq \alpha) \quad (3.3)$$

onde α é o quantil $1-\alpha$ de $H(\alpha)$. Para o desconhecido, o ponto crítico α é determinado a partir de $H(\alpha)$, sendo igual a:

$$H(\alpha) = P(\hat{\theta} - \hat{\theta}^* \leq \alpha) \quad (3.4)$$

onde α é o quantil $1-\alpha$ de $H(\alpha)$. Estas expressões são consideradas exatas porque verificam:

3 - INTERVALOS DE CONFIANÇA *BOOTSTRAP*

Neste capítulo será abordado o problema da determinação de Intervalos de confiança aproximados para o parâmetro θ recorrendo às técnicas *Bootstrap*.

A dedução de um intervalo de confiança exacto para θ pode ser facilmente prosseguida conhecendo uma das distribuições, $H(x)$ ou $K(x)$ conforme σ seja conhecido ou desconhecido, com,

$$H(x) = P\{\sqrt{n}(\hat{\theta} - \theta)/\sigma \leq x\} , \quad (3.1)$$

$$K(x) = P\{\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma} \leq x\} ,$$

em que $\hat{\theta}$ é o estimador da M.V. de θ com variância assintótica σ^2/n e $\hat{\sigma}$ representa uma estimativa de σ obtida geralmente com base na Informação de Fisher.

Os extremos do Intervalo de Confiança para θ são definidos com base no ponto crítico $\hat{\theta}(\alpha)$ que verifica $P\{\theta \leq \hat{\theta}(\alpha)\} = \alpha$. Com σ conhecido, o ponto crítico é dado por,

$$\hat{\theta}_{ord}(\alpha) = \hat{\theta} - n^{-1/2} \sigma x_{1-\alpha} , \quad (3.2)$$

sendo $x_{1-\alpha}$ o quantil $1-\alpha$ de $H(x)$. Para σ desconhecido, o ponto crítico é deduzido a partir de $K(x)$, sendo igual a,

$$\hat{\theta}_{stud}(\alpha) = \hat{\theta} - n^{-1/2} \hat{\sigma} y_{1-\alpha} , \quad (3.3)$$

com $y_{1-\alpha}$ o quantil $1-\alpha$ de $K(x)$. Estes extremos são considerados exactos porque verificam,

$$P\{\theta \leq \hat{\theta}_{ord}(\alpha)\} = P\{\theta \leq \hat{\theta}_{stud}(\alpha)\} = \alpha .$$

No entanto na grande maioria das situações não é possível o conhecimento de I.C. exactos. Neste contexto, a determinação de I.C. aproximados assume grande relevância nos problemas de inferência. Um I.C. aproximado muito utilizado é o intervalo *standard* a $(1-2\alpha) \times 100\%$ de confiança,

$$\theta \in (\hat{\theta} + n^{-1/2} \hat{\sigma} z_{\alpha} , \hat{\theta} + n^{-1/2} \hat{\sigma} z_{1-\alpha}) . \quad (3.4)$$

A determinação de (3.4) tem como pressuposto que a distribuição de $\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}$, $K(x)$, é assintoticamente uma Normal estandardizada, por outras palavras,

$$K(x) = \Phi(x) + R(x) , \quad (3.5)$$

em que $R(x)$ é negligenciável para n suficientemente grande. Expandindo $R(x)$ em série é possível encontrar aproximações a $K(x)$ mais precisas, como a aproximação dada pela expansão de Edgeworth.

Suponha-se que $H(x)$ e $K(x)$ verificam uma expansão de Edgeworth do tipo,

$$H(x) = \Phi(x) + \sum_{i=1}^v n^{-i/2} p_i(x) \phi(x) + O(n^{-(v+1)/2}) \quad (3.6) ,$$

$$K(x) = \Phi(x) + \sum_{i=1}^v n^{-i/2} q_i(x) \phi(x) + O(n^{-(v+1)/2}) ,$$

com $p_i(x)$ e $q_i(x)$ polinómios de grau $3i-1$, sendo os polinómios de índice inferior par/ímpar funções ímpares/pares. Os quantis de ordem α das referidas distribuições, respectivamente x_{α} e y_{α} são dados pela expansão (inversa) de Cornish Fisher,

$$x_a = H^{-1}(\alpha) = z_a + \sum_{i=1}^{\nu} n^{-i/2} p_{i1}(z_a) + O(n^{-(\nu+1)/2}) , \quad (3.7)$$

$$y_a = K^{-1}(\alpha) = z_a + \sum_{i=1}^{\nu} n^{-i/2} q_{i1}(z_a) + O(n^{-(\nu+1)/2}) ,$$

onde $p_{i1}(x)$ e $q_{i1}(x)$ são polinómios de grau $i+1$ e os polinómios de índice inferior par/ímpar são respectivamente funções ímpares/pares. A partir de (3.7) é possível deduzir uma expansão dos extremos exactos¹ do I.C. para θ .

$$\hat{\theta}_{ord}(\alpha) = \hat{\theta} + n^{-1/2} \sigma \{ z_a + n^{-1/2} p_1(z_a) + n^{-1} p_{21}(z_a) \} + O(n^{-2}) , \quad (3.8)$$

$$\hat{\theta}_{stud}(\alpha) = \hat{\theta} + n^{-1/2} \hat{\sigma} \{ z_a + n^{-1/2} q_1(z_a) + n^{-1} q_{21}(z_a) \} + O_p(n^{-2}) ,$$

sabendo que $p_{11}(x) = -p_1(x)$ e $q_{11}(x) = -q_1(x)$.

Uma forma de avaliar a qualidade dos I.C. aproximados consiste em comparar as expansões assintóticas dos respectivos extremos com as expansões dos extremos correctos patentes em (3.8). Concretizando para o intervalo *standard*, nota-se que o extremo $\hat{\theta} + n^{-1/2} \hat{\sigma} z_a$ denota um erro da ordem $1/n$ relativamente ao extremo exacto patente na expressão inferior de (3.8), o que leva a considerar este intervalo correcto de primeira ordem.

¹ Peter Hall(1988) p. 944

É interessante comparar, nos I.C. bilaterais, a amplitude dos intervalos aproximados com a amplitude dos intervalos exactos. A amplitude do I.C. exacto $[\hat{\theta}_{ord}(\alpha), \hat{\theta}_{ord}(1-\alpha)]$ é,

$$\begin{aligned} A_{ord}(1-2\alpha) &= \hat{\theta}_{ord}(1-\alpha) - \hat{\theta}_{ord}(\alpha) \\ &= 2n^{-1/2} \sigma \{z_{1-\alpha} + n^{-1} p_{21}(z_{1-\alpha})\} + O(n^{-5/2}), \end{aligned} \quad (3.9)$$

enquanto que a amplitude do I.C. $[\hat{\theta}_{stud}(\alpha), \hat{\theta}_{stud}(1-\alpha)]$ tem como expressão,

$$\begin{aligned} A_{stud}(1-2\alpha) &= \hat{\theta}_{stud}(1-\alpha) - \hat{\theta}_{stud}(\alpha) \\ &= 2n^{-1/2} \hat{\sigma} \{z_{1-\alpha} + n^{-1} q_{21}(z_{1-\alpha})\} + O_p(n^{-5/2}). \end{aligned} \quad (3.10)$$

O intervalo bilateral *standard* tem amplitude,

$$A_{stand}(1-2\alpha) = \hat{\theta}_{stand}(1-\alpha) - \hat{\theta}_{stand}(\alpha) = 2n^{-1/2} \hat{\sigma} z_{1-\alpha}, \quad (3.11)$$

donde $|A_{stand}(1-2\alpha) - A_{stud}(1-2\alpha)| = O(n^{-3/2})$. Isto significa que ao aproximar-se um dado I.C. pelo intervalo *standard* comete-se um erro na amplitude da ordem de $n^{-3/2}$.

A qualidade dos I.C. aproximados não se reflecte apenas na forma como se ajustam os seus extremos aos verdadeiros extremos. Outro aspecto interessante consiste em avaliar as respectivas coberturas reais pois, dado que são intervalos aproximados, estas não coincidem necessariamente com a cobertura nominal. O intervalo *standard* unilateral a $\alpha \times 100\%$, $[-\infty, \hat{\theta} + n^{-1/2} \hat{\sigma} z_\alpha]$ tem cobertura¹

$$\begin{aligned} \pi_{stand}(\alpha) &= \alpha - n^{-1/2} q_1(z_\alpha) \phi(z_\alpha) + n^{-1} q_2(z_\alpha) \phi(z_\alpha) + O(n^{-3/2}) \\ &= \alpha + O(n^{-1/2}), \end{aligned} \quad (3.12)$$

¹ Peter Hall(1988) p. 944 .

onde $q_1(x)$ e $q_2(x)$ são os polinómios da E.E. (3.6). Por seu turno, a probabilidade do intervalo bilateral a $(1-2\alpha)\times 100\%$, $[\hat{\theta} + n^{-1/2} \hat{\sigma} z_{1-\alpha}, \hat{\theta} + n^{-1/2} \hat{\sigma} z_{1-\alpha}]$ é,¹

$$\begin{aligned} \pi_{stand}(1-\alpha) - \pi_{stand}(\alpha) &= 1 - 2\alpha + 2n^{-1} q_2(z_{1-\alpha}) \phi(z_{1-\alpha}) + O(n^{-2}) \\ &= 1 - 2\alpha + O(n^{-1}) \end{aligned} \quad (3.13)$$

Estas expressões mostram que o erro de cobertura dos I.C. *standard* unilaterais é de ordem $n^{-1/2}$, enquanto que nos bilaterais é de ordem n^{-1} .

O desafio que o Bootstrap enfrenta consiste em procurar determinar I.C. aproximados que tal como o *standard* se apliquem de forma automática mas em contrapartida detenham qualidades superiores de aproximação. Estas reflectem-se por um lado na obtenção de extremos correctos de ordem superior (é difícil avançar para além da correcção de segunda ordem porque normalmente os extremos aproximados envolvem estimativas dos polinómios $p_i(x)$ e $q_i(x)$ a partir de $i=1$); e por outro lado, na obtenção de uma probabilidade de cobertura real mais próxima da cobertura nominal, não esquecendo também a justeza da aproximação da amplitude do intervalo. Estes aspectos vão estar patentes na análise dos vários métodos considerados de construção de I.C. bootstrap - percentis, correcção do enviesamento (BC), correcção do enviesamento e aceleração (BCa), e o *bootstrap-t*.

¹ Peter Hall(1988) p.948

3.1 MÉTODO DOS PERCENTIS

O método dos percentis consiste em tomar os percentis da distribuição bootstrap de $\hat{\theta}$ como os extremos do I.C. para θ . Considere-se $\hat{\theta}_p(\alpha)$ como o quantil de ordem α de $\hat{C}(s)$.

$$\hat{\theta}_p(\alpha) = \hat{C}^{-1}(\alpha) . \quad (3.14)$$

O I.C. para θ unilateral a $\alpha \times 100\%$ é,

$$\theta \in [-\infty , \hat{\theta}_p(\alpha)] ,$$

enquanto que o intervalo bilateral a $(1-2\alpha) \times 100\%$ é,

$$\theta \in [\hat{\theta}_p(\alpha) , \hat{\theta}_p(1-\alpha)] .$$

Sendo denotadamente simples e intuitivo, este método tem alguns fundamentos teóricos válidos. Por um lado, os I.C. apresentados podem ter uma interpretação Bayesiana e, por outro lado, podem ser deduzidos supondo uma dada transformação normalizadora. Estas interpretações são desenvolvidas nos parágrafos seguintes.

3.1.1 - Justificação Bayesiana do método dos percentis

Os I.C. obtidos com o método dos percentis podem ser considerados muito próximos de I.C. Bayesianos quando se verificam determinadas condições específicas.

Para explicitar estas condições considere-se que o universo do problema é constituído por valores discretos¹, $X = \{ 1, \dots, l, \dots, L \}$

¹ Esta situação pode ser facilmente generalizada para o caso contínuo, veja-se Efron(1982a) p. 81.

e defina-se $f_l = P\{X=l\}$ e $\hat{f}_l = \#\{x_i=l\}/n$. Quando a distribuição *a priori* para \underline{f} é uma distribuição de Dirichlet simétrica com parâmetro k , $\underline{f} \sim D_u(k\underline{e})$ com $\underline{e}=[1\dots 1]^t$, a respectiva distribuição *a posteriori* tendo sido observado $\underline{\hat{f}} = [\hat{f}_1 \dots \hat{f}_L]^t$ é uma $D_u(k\underline{e} + n\underline{\hat{f}})$. Supondo que prevalece a ignorância *a priori* com $k \rightarrow 0$ então a distribuição *a posteriori* resume-se a $\underline{f}|\underline{\hat{f}} \sim D_u(n\underline{\hat{f}})$

Considerando o problema na perspectiva *bootstrap* supõem-se que, à amostra *bootstrap* corresponde o vector de probabilidades $\underline{f}^* = [f_1^* \dots f_L^*]^t$ onde nf_i^* representa o número de ocorrências da categoria i de entre L categorias possíveis em que cada categoria j tem probabilidade \hat{f}_j , ($j=1, \dots, L$), numa sucessão de n provas consecutivas. Desta forma, $\underline{f}^*|\underline{\hat{f}} \sim Mult_L(n, \underline{\hat{f}})/n$. Esta distribuição é muito semelhante à $D_u(n\underline{\hat{f}})$, nomeadamente quando $\theta(F)$ é uma função razoavelmente suave de F .

Nas condições enunciadas o I.C. central obtido com o método dos percentis constitui uma boa aproximação do I.C. bayesiano. No entanto, os pressupostos assumidos nesta formalização são algo questionáveis e restrictivos. As considerações *a priori* de uma distribuição Dirichlet e a prevalência de ignorância podem não ser muito adequadas e realistas: é bem possível que a distribuição *a priori* contenha outros valores para além dos valores observados. Em contrapartida é verosímil que se detenha uma informação prévia que permita considerar irrelevantes algumas observações no processo de reamostragem. Estes aspectos denotam uma certa fragilidade do método dos percentis na sua óptica bayesiana.

3.1.2 - Abordagem com base numa transformação normalizadora

Em muitos problemas existe a possibilidade de definir uma transformação monótona que aplica, pelo menos com grande aproximação, a estatística em causa numa variável pivô Normal¹. Este procedimento tem inerente a vantagem de deslocar o problema de uma distribuição complicada para a Normal, simplificando desta forma o processo de inferência. Efron utilizou este artifício para motivar o método dos percentis.

Este método tem como ponto de partida a existência de uma transformação monótona normalizadora, g , tal que, para,

$$\phi = g(\theta) \quad , \quad \hat{\phi} = g(\hat{\theta}) \quad , \quad \hat{\phi}^* = g(\hat{\theta}^*) \quad , \quad (3.15a)$$

se verifica,

$$\frac{\hat{\phi} - \phi}{\tau} \sim \mathcal{N}(0, 1) \quad , \quad \frac{\hat{\phi}^* - \hat{\phi}}{\tau} \sim \mathcal{N}(0, 1) \quad , \quad (3.15)$$

sendo τ uma constante positiva.

Conhecendo g é fácil determinar o I.C. para θ . Calcula-se primeiro o I.C. para ϕ que, de acordo com (3.15) tem extremos do tipo $\hat{\phi} \pm \tau z_\alpha$, e depois, utilizando a transformação inversa de g , aplicam-se estes extremos em valores de θ obtendo-se $g^{-1}(\hat{\phi} \pm \tau z_\alpha)$.

O *bootstrap* efectua este processo automaticamente sem haver necessidade do conhecimento de g , apenas há que postular a sua existência.

¹ Como exemplo, recorde-se a muito conhecida transformação do coeficiente de correlação, \tanh^{-1} . Sobre este assunto veja-se Efron (1982b).

Análise-se com mais detalhe o procedimento referido. De acordo com (3.15) verifica-se,

$$\begin{aligned}
 P\{\hat{\phi} < \phi + \tau z_{\alpha}\} &= \alpha & \Leftrightarrow & P_{*}\{\hat{\phi}^{*} < \hat{\phi} + \tau z_{\alpha}\} = \alpha \\
 & & \Leftrightarrow & P_{*}\{g(\hat{\theta}^{*}) < \hat{\phi} + \tau z_{\alpha}\} = \alpha \\
 & & \Leftrightarrow & P_{*}\{\hat{\theta}^{*} < g^{-1}(\hat{\phi} + \tau z_{\alpha})\} = \alpha \\
 & & \Leftrightarrow & \hat{G}(g^{-1}(\hat{\phi} + \tau z_{\alpha})) = \alpha
 \end{aligned}
 \tag{3.16}$$

Perante (3.16) é possível concluir que, admitindo a hipótese da existência de uma transformação monótona g verificando (3.15), $\hat{\theta}_p(\alpha) = g^{-1}(\hat{\phi} + \tau z_{\alpha})$ e desta forma, $[-\infty, \hat{\theta}_p(\alpha)]$ é um intervalo exacto na medida que a sua cobertura é α . De igual modo, o I.C. bilateral $[\hat{\theta}_p(\alpha), \hat{\theta}_p(1-\alpha)]$ tem cobertura $1-2\alpha$.

As hipóteses fixadas em (3.15) pressupõem que,

$$P\{\hat{\phi} \leq \phi\} = 0.5 \Leftrightarrow P\{\hat{\theta} \leq \theta\} = P_{*}\{\hat{\theta}^{*} \leq \hat{\theta}\} = 0.5 \quad . \tag{3.17}$$

No entanto, frequentemente o estimador $\hat{\theta}$ é enviesado no sentido que $P\{\hat{\theta} \leq \theta\} \neq 0.5$. Nestas condições não existe nenhuma transformação g monótona que verifique (3.15) abalando a validade do método dos percentis. Efron realizou vários estudos de simulação tendo concluído existir uma certa incapacidade do método em captar fielmente o enviesamento e a amplitude do intervalo exacto.¹ Esta questão motivou o autor a introduzir uma correcção que contemplasse o possível enviesamento de $\hat{\theta}$ dando origem ao método dos percentis com correcção do enviesamento (BC).

¹ Veja-se Efron(1982a) p. 79 ; pp. 85 e 86 ; pp. 88, 89 e 90 ; e Efron(1983) p. 41.

Do exposto é possível concluir que o método dos percentis é um tanto vulnerável. Os fundamentos do método associados ao pressuposto da existência de transformação normalizadora contemplam uma situação muito particular (por exemplo, deixam de ser válidos quando o estimador é enviesado). De igual modo, na interpretação bayesiana, para que $\hat{G}(s)$ represente aproximadamente a distribuição *a posteriori* de θ , há que impôr considerações *a priori* muito restritas e pouco realistas (por exemplo a distribuição *a priori* tem de ser uma $D_{ik}(k \geq 0)$ com $k \rightarrow 0$). Dados estes factos é de supôr que, perante um problema concreto, os I.C. obtidos com o método dos percentis podem não aproximar de forma muito correcta os verdadeiros I.C.. Através das expansões de Edgeworth e de Cornish-Fisher é possível compreender de forma mais precisa a relação entre a aproximação dada pelo método dos percentis e os I.C. exactos.

3.1.3 - Propriedades do método dos percentis

As versões *bootstrap* das distribuições $H(x)$ e $K(x)$ são respectivamente,

$$\hat{H}(x) = P\{\sqrt{n}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma} \leq x \mid \hat{X}\}, \quad (3.18)$$

$$\hat{K}(x) = P\{\sqrt{n}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^* \leq x \mid \hat{X}\},$$

(\hat{X} representa a amostra observada).

Supondo que estas distribuições possuem expansão de Edgeworth, vem,

$$\hat{H}(x) = \Phi(x) + \sum_{i=1}^p n^{-1/2} \hat{p}_i(x) \phi(x) + O_p(n^{-(p+1)/2}) \quad , \quad (3.19)$$

$$\hat{K}(x) = \Phi(x) + \sum_{i=1}^p n^{-1/2} \hat{q}_i(x) \phi(x) + O_p(n^{-(p+1)/2}) \quad ,$$

onde $\hat{p}_i(x)$ e $\hat{q}_i(x)$ são polinómios que resultam de substituir em $p_i(x)$ e $q_i(x)$ as quantidades desconhecidas por estimativas *bootstrap*.

O ponto crítico deduzido através do método dos percentis, $\hat{\theta}_p(\alpha)$, corresponde à estimativa *bootstrap* do ponto teórico que Hall designa por,

$$\hat{\theta}_{back}(\alpha) = \hat{\theta} + n^{-1/2} \hat{\sigma} x_\alpha \quad . \quad (3.20)$$

Essa estimativa obtém-se substituindo x_α pela sua estimativa *bootstrap* \hat{x}_α , o quantil de ordem α de $\hat{H}(x)$!

Na construção de $\hat{\theta}_{back}$ estão inerentes dois tipos de erros. Por um lado, ao utilizar-se na sua definição $\hat{\sigma}$ deveria considerar-se um quantil de $K(x)$ e não de $H(x)$ (compare-se com (3.3); por outro lado, tal quantil devia ser obtido como o simétrico do quantil $1-\alpha$ e não o quantil de ordem α (compare-se com (3.2)). Estes erros vão repercutir-se no I.C. baseado em $\hat{\theta}_p(\alpha)$, dado que este é a versão *bootstrap* de $\hat{\theta}_{back}(\alpha)$, sendo responsáveis pela falta de precisão do método dos percentis.

1 Veja-se que,

$$P\{\sqrt{n}(\hat{\theta}^* - \hat{\theta})/\hat{\sigma} \leq \hat{x}_\alpha\} = \alpha \Leftrightarrow P\{\hat{\theta}^* \leq \hat{\theta} + \hat{\sigma} n^{-1/2} \hat{x}_\alpha\} = \hat{G}(\hat{\theta} + \hat{\sigma} n^{-1/2} \hat{x}_\alpha) = \alpha$$

onde $\hat{\theta} + \hat{\sigma} n^{-1/2} \hat{x}_\alpha$ é o quantil de ordem α de $\hat{G}(s)$.

O I.C. unilateral aproximado determinado a partir do ponto teórico $\hat{\theta}_{back}(\alpha)$ tem probabilidade de cobertura igual a,¹

$$\begin{aligned}\pi_{back}(\alpha) &= P\{\theta \leq \hat{\theta}_{back}(\alpha)\} \\ &= \alpha - n^{-1/2} \{p_1(z_\alpha) + q_1(z_\alpha)\} \phi(z_\alpha) + O(n^{-1}) \\ &= \alpha + O(n^{-1/2}) \quad .\end{aligned}\quad (3.21)$$

A probabilidade de cobertura dos I.C. obtidos com o método dos percentis, respectivamente intervalos unilaterais e bilaterais, é dada por,²

$$\begin{aligned}\pi_p(\alpha) &= P\{\theta \leq \hat{\theta}_p(\alpha)\} \\ &= \alpha - n^{-1/2} \{\hat{p}_1(z_\alpha) + \hat{q}_1(z_\alpha)\} \phi(z_\alpha) + O_p(n^{-1}) \\ &= \alpha + O_p(n^{-1/2}) \quad ,\end{aligned}\quad (3.22)$$

$$\begin{aligned}\pi_p(1-\alpha) - \pi_p(\alpha) &= (1-2\alpha) - 2n^{-1} \left\{ -\frac{1}{2} \hat{p}_1(z_{1-\alpha})^2 z_{1-\alpha} \right. \\ &\quad \left. - \hat{p}_1(z_{1-\alpha}) [q'_1(z_{1-\alpha}) - q_1(z_{1-\alpha}) z_{1-\alpha}] \right. \\ &\quad \left. - q_2(z_{1-\alpha}) - \hat{p}_{21}(z_{1-\alpha}) + u z_{1-\alpha} \right\} \phi(z_{1-\alpha}) + O_p(n^{-2}) \\ &= (1-2\alpha) + O_p(n^{-1}) \quad ,\end{aligned}\quad (3.23)$$

onde u é uma constante que satisfaz,

$$E\{[\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}][\sqrt{n}(\hat{p}_1(z_\alpha) - p_1(z_\alpha))]\} = u + O(n^{-1}) \quad .$$

É de notar que a grandeza do erro de cobertura dos I.C. produzidos com este método é idêntica à dos I.C. *standard*, nomeadamente para os intervalos unilaterais o erro é da ordem $n^{-1/2}$ e para os bilaterais é da ordem n^{-1} , embora para os intervalos *bootstrap* a ordem do erro seja apenas estabelecida em probabilidade.

1 Hall(1988) p. 934.

2 Veja-se Peter Hall(1988) p. 947 (intervalos unilaterais) e p. 948 (intervalos bilaterais).

Por sua vez, o ponto crítico $\hat{\theta}_p(\alpha)$ verifica a expansão,¹

$$\hat{\theta}_p(\alpha) = \hat{\theta} + n^{-1/2} \hat{\sigma} \{ z_\alpha - n^{-1/2} \hat{p}_1(z_\alpha) + n^{-1} \hat{p}_{21}(z_\alpha) \} + O_p(n^{-2}) \quad (3.24)$$

sendo $\hat{p}_{11} = -\hat{p}_1$, \hat{p}_{11} e \hat{p}_{21} os polinómios que aparecem na expansão (inversa) de Cornish-Fisher de \hat{x}_α obtidos substituindo nos p_i , (dados em (3.7)) as quantidades desconhecidas por estimativas *bootstrap*.

Comparando (3.24) com a expansão do extremo exacto $\hat{\theta}_{stud}(\alpha)$ dada por (3.8) nota-se que $\hat{\theta}_p(\alpha)$ difere deste por substituir os polinómios q_1 e q_{21} por estimativas dos polinómios p_1 e p_{21} , isto é, por estimativas obtidas a partir de $H(x)$ quando deveria ser de $K(x)$. Quando $H(x)$ é diferente de $K(x)$, $\hat{\theta}_p(\alpha)$ é correcto de primeira ordem visto que, tal como o extremo do I.C. *standard*, só aproxima correctamente até ao termo $n^{-1/2} \hat{\sigma} z_\alpha$, verificando-se $|\hat{\theta}_p - \hat{\theta}_{stud}| = O_p(n^{-1})$. No entanto, $\hat{\theta}_p(\alpha)$ entra em conta de certa forma com os termos assintóticos de ordem superior que muitas vezes são os responsáveis pela assimetria dos I.C. em relação a $\hat{\theta}$ (estimativa da M.V.), donde em certas situações $\hat{\theta}_p(\alpha)$ é capaz de determinar melhores I.C. que os *standard* por poder captar melhor essa assimetria.

A amplitude do I.C. $[\hat{\theta}_p(\alpha), \hat{\theta}_p(1-\alpha)]$ é igual a,

$$\begin{aligned} A_p(1-2\alpha) &= \hat{\theta}_p(1-\alpha) - \hat{\theta}_p(\alpha) \\ &= 2n^{-1/2} \hat{\sigma} \{ z_{1-\alpha} + n^{-1} \hat{p}_{21}(z_{1-\alpha}) \} + O_p(n^{-5/2}) \\ &= 2n^{-1/2} \hat{\sigma} z_{1-\alpha} + O_p(n^{-3/2}) \quad (3.25) \end{aligned}$$

¹ Peter Hall(1988) p. 943.

Este intervalo apresenta um erro de amplitude da ordem $n^{-3/2}$ tal como o intervalo *standard* (embora o primeiro seja em probabilidade). No entanto, estes intervalos são centrados em valores distintos, pelos motivos atrás referidos.

Como foi apontado, $\hat{\theta}_p(\alpha)$ é a estimativa *bootstrap* de $\hat{\theta}_{back}(\alpha)$, extremo teórico não exacto do I.C. para θ . A inexactidão é provocada por $\hat{\theta}_{back}(\alpha)$ ser definido a partir de x_α em vez de $-y_{1-\alpha}$. Um ponto crítico mais exacto pode ser determinado através da manipulação de α de forma a obter um outro quantil x_β mais próximo de $-y_{1-\alpha}$. A repercussão *bootstrap* desta medida consiste em considerar como extremo do I.C. um outro quantil, de ordem β , de $\hat{G}(s)$ em vez de $\hat{\theta}_p(\alpha)$ de forma a que os I.C. *bootstrap* $[-\infty, \hat{\theta}_p(\beta)]$ e $[\hat{\theta}_p(\beta), \hat{\theta}_p(1-\beta)]$ tenham uma cobertura mais próxima da cobertura nominal e verifiquem uma melhor aproximação dos extremos exactos. É com base neste procedimento que funcionam as correcções introduzidas por Efron ao método dos percentis - primeiro a correcção do enviesamento e posteriormente a correcção da aceleração.

3.2 - MÉTODO COM CORRECÇÃO DO ENVIESAMENTO

A correcção do enviesamento é feita com base na constante z_0 definida por,

$$z_0 = \Phi^{-1}[\hat{G}(\hat{\theta})] , \quad (3.26)$$

determinando um valor para β de acordo com,

$$\beta = \Phi(2z_0 + z_\alpha) . \quad (3.27)$$

O intervalo de confiança unilateral a $\alpha \times 100\%$ obtido com este método tem como extremo,

$$\hat{\theta}_{bc}(\alpha) = \hat{\theta}_p(\beta) = \hat{G}^{-1}[\Phi(2z_0 + z_\alpha)] , \quad (3.28)$$

enquanto que o I.C. bilateral a $(1-2\alpha) \times 100\%$ é definido por,

$$[\hat{G}^{-1}[\Phi(2z_0 + z_\alpha)] , \hat{G}^{-1}[\Phi(2z_0 - z_\alpha)]] . \quad (3.29)$$

É de notar que quando $\hat{G}(\hat{\theta}) = 0.5$, situação em que $\hat{\theta}$ é centrado na mediana, z_0 é igual a zero e como consequência $\beta = \alpha$, coincidindo o I.C. BC com o intervalo obtido com o método dos percentis.

3.2.1 - Abordagem com base numa transformação normalizadora

O método dos percentis produz I.C. para θ exactos quando existe uma transformação normalizadora que verifica (3.15). Tal não acontece quando $z_0 \neq 0$. Neste caso, a transformação monótona g , com,

$$\phi = g(\theta) \quad , \quad \hat{\phi} = g(\hat{\theta}) \quad \text{e} \quad \hat{\phi}^* = g(\hat{\theta}^*) \quad , \quad (3.30a)$$

tem de verificar,

$$\frac{\hat{\phi} - \phi}{\tau} \sim \mathcal{N}(-z_0, 1) \quad \text{e} \quad \frac{\hat{\phi}^* - \hat{\phi}}{\tau} \sim \mathcal{N}(-z_0, 1) \quad , \quad (3.30)$$

dados τ e z_0 constantes.

Recorrendo às hipóteses apresentadas, a justificação da expressão (3.26) que define z_0 é muito simples. Sabe-se que,

$$P.\{\hat{\theta}^* \leq \hat{\theta}\} = P.\{\hat{\phi}^* \leq \hat{\phi}\} = P.\left\{\frac{\hat{\phi}^* - \hat{\phi}}{\tau} + z_0 \leq z_0\right\} \quad ,$$

donde se obtém,

$$P.\{Z \leq z_0\} = \hat{G}(\hat{\theta}) \quad ,$$

com Z uma variável Normal estandardizada, e como tal $z_0 = \Phi^{-1}[\hat{G}(\hat{\theta})]$.

Para deduzir (3.29) apela-se a um raciocínio coincidente com o apresentado no método anterior. A partir da variável pivô $\hat{\phi} - \phi$ determina-se o I.C. para ϕ que ao nível $(1-2\alpha) \times 100\%$ coincide com

$$\phi \in [\hat{\phi} + z_0\tau + z_\alpha\tau \quad , \quad \hat{\phi} + z_0\tau - z_\alpha\tau] \quad ,$$

e aplicando a este intervalo a transformação inversa de g , obtém-se o I.C. para θ .

$$\theta \in [g^{-1}(\hat{\phi} + z_0\tau + z_\alpha\tau) \quad , \quad g^{-1}(\hat{\phi} + z_0\tau - z_\alpha\tau)] \quad .$$

Como $\hat{\phi} - \phi$ e $\hat{\phi}^* - \hat{\phi}$ têm distribuição idêntica, verifica-se,
 $P\{\hat{\phi} \leq \phi + z_0 \tau + z_\alpha \tau\} = P\{\hat{\phi}^* \leq \hat{\phi} + z_0 \tau + z_\alpha \tau\} = \Phi(2z_0 + z_\alpha)$,
 dado que a partir de (3.30) se tem $(\hat{\phi}^* - \hat{\phi})/\tau + z_0 \sim \mathcal{N}(0,1)$. Por
 seu turno,

$$\begin{aligned} P\{\hat{\phi}^* \leq \hat{\phi} + z_0 \tau + z_\alpha \tau\} &= P\{g(\hat{\theta}^*) \leq \hat{\phi} + z_0 \tau + z_\alpha \tau\} \\ &= P\{\hat{\theta}^* \leq g^{-1}(\hat{\phi} + z_0 \tau + z_\alpha \tau)\} \\ &= \Phi(2z_0 + z_\alpha) , \end{aligned}$$

donde se conclui que $g^{-1}(\hat{\phi} + z_0 \tau + z_\alpha)$ corresponde a $\hat{G}^{-1}[\Phi(2z_0 + z_\alpha)]$,
 coincidindo com o extremo inferior do I.C. (3.29). O mesmo
 raciocínio é válido para o extremo superior.

O I.C. BC pode ser obtido directamente a partir da distribuição
bootstrap de $\hat{\theta}$ sem conhecimento da transformação g - basta
 postular a sua existência. Verificando-se (3.30) este intervalo
 tem probabilidade de cobertura real coincidente com a cobertura
 nominal e, deste modo, é um intervalo exacto.

3.2.2 - Propriedades do BC

O ponto crítico $\hat{\theta}_{bc}(\alpha)$ corresponde à estimativa *bootstrap*
 de,

$$\hat{\theta}_{back}(\beta) = \hat{\theta} + n^{-1/2} \hat{\sigma} x_\beta , \quad (3.31)$$

com, $\beta = \Phi\{2m + z_\alpha\}$, sendo $m = \Phi^{-1}[G(\hat{\theta})]$, $G(x) = P\{\hat{\theta} \leq x\}$. Esta
 estimativa obtém-se substituindo x_β por \hat{x}_β , o $\beta \times 100$ -ésimo
 percentil de $\hat{H}(x)$, considerando para β o valor estimado dado
 em (3.27).

A probabilidade de cobertura do I.C. unilateral a $\alpha \times 100\%$ com base em $\hat{\theta}_{back}(\alpha)$ é, de acordo com (3.21), proporcional em grandes amostras ao polinómio quadrático $p_1(z_\alpha) + q_1(z_\alpha)$. A correcção do enviesamento vai eliminar a constante deste polinómio (igual a $2p_1(0)$ dado que $q_1(0) = p_1(0)$ pois $H(0) = K(0)$), sendo a cobertura do I.C. unilateral a $\alpha \times 100\%$ com extremo $\hat{\theta}_{back}(\beta)$ igual a,¹

$$\begin{aligned}\pi_{back}(\beta) &= P\{\theta \leq \hat{\theta}_{back}(\beta)\} \\ &= \alpha - n^{-1/2} \{p_1(z_\alpha) + q_1(z_\alpha) - 2p_1(0)\} \phi(z_\alpha) + O(n^{-1}) \\ &= \alpha + O(n^{-1/2}) .\end{aligned}\quad (3.32)$$

A probabilidade de cobertura do I.C. unilateral a $\alpha \times 100\%$, $[-\infty, \hat{\theta}_{bc}(\alpha)]$ e do intervalo bilateral a $(1-2\alpha) \times 100\%$, $[\hat{\theta}_{bc}(\alpha), \hat{\theta}_{bc}(1-\alpha)]$, são dadas respectivamente por $\pi_{bc}(\alpha)$ e $\pi_{bc}(1-\alpha) - \pi_{bc}(\alpha)$ com,²

$$\begin{aligned}\pi_{bc}(\alpha) &= P\{\theta \leq \hat{\theta}_{bc}(\alpha)\} \\ &= \alpha - n^{-1/2} \{\hat{p}_1(z_\alpha) + \hat{q}_1(z_\alpha) - 2\hat{p}_1(0)\} \phi(z_\alpha) + O_p(n^{-1}) \\ &= \alpha + O_p(n^{-1/2}) ,\end{aligned}\quad (3.33)$$

$$\begin{aligned}\pi_{bc}(1-\alpha) - \pi_{bc}(\alpha) &= (1-2\alpha) - 2n^{-1} \left\{ \hat{p}_1(0) - \frac{1}{2} \hat{p}_1(z_{1-\alpha})^2 z_{1-\alpha} \right. \\ &\quad + [2\hat{p}_1(0) - \hat{p}_1(z_{1-\alpha})][q'_1(z_{1-\alpha}) \\ &\quad - q_1(z_{1-\alpha}) z_{1-\alpha}] - q_2(z_{1-\alpha}) - \hat{p}_{21}(z_{1-\alpha}) \\ &\quad \left. + 2p'_1(z_{1-\alpha}) \hat{p}_1(0) + u z_{1-\alpha} \right\} \phi(z_{1-\alpha}) + O_p(n^{-2}) \\ &= (1-2\alpha) + O_p(n^{-1}) ,\end{aligned}\quad (3.34)$$

1 Peter Hall(1988) p.934

2 Peter Hall(1988) p.943 (I.C. unilaterais) e p.948 (I.C. bilaterais).

onde u é uma constante que satisfaz,

$$E \left\{ \left[\sqrt{n} \frac{(\hat{\theta} - \theta)}{\hat{\sigma}} \right] [\sqrt{n} (2\hat{p}_1(0) - 2p_1(0) + \hat{p}_1(z_\alpha) - p_1(z_\alpha))] \right\} = u + O(n^{-1}) .$$

Nestas expressões está patente que a correcção do enviesamento não diminui a ordem de grandeza do erro da probabilidade de cobertura dos I.C. patente nos intervalos obtidos com os métodos *standard* e dos percentis.

O ponto crítico $\hat{\theta}_{bc}(\alpha)$ verifica a expansão,¹

$$\begin{aligned} \hat{\theta}_{bc}(\alpha) = & \hat{\theta} + n^{-1/2} \hat{\sigma} \{ z_\alpha - n^{-1/2} [\hat{p}_1(z_\alpha) - 2\hat{p}_1(0)] \\ & + n^{-1} [\hat{p}_{21}(z_\alpha) - 2\hat{p}'_1(z_\alpha)\hat{p}_1(0)] \} + O_p(n^{-2}) \end{aligned} \quad (3.35)$$

É de notar que sendo $\hat{p}_1(0)$ nulo então $\hat{\theta}_{bc}(\alpha) = \hat{\theta}_p(\alpha)$ e o intervalo obtido com o método dos percentis coincide com o intervalo BC. Este resultado é coerente com o facto de $p_1(0)$ traduzir o enviesamento da estatística pois,

$$m = \Phi^{-1} \{ H(0) \} = n^{-1/2} p_1(0) + O(n^{-1}) ,$$

logo quando $p_1(0) = 0$ vem $m = 0$ e por consequência $\beta = \alpha$. Comparando a expansão de $\hat{\theta}_{bc}(\alpha)$ com a expansão do extremo exacto $\hat{\theta}_{stud}(\alpha)$ expressa em (3.8) conclui-se que $\hat{\theta}_{bc}(\alpha)$ é correcto de primeira ordem (tal como os outros pontos críticos analisados anteriormente), não excluindo no entanto que para alguns problemas $\hat{\theta}_{bc}(\alpha)$ seja correcto de segunda ordem. São exemplo os problemas referidos em Efron(1985) onde a população tem distribuição Normal k -dimensional com um dado vector de médias e matriz de covariâncias igual à Identidade, e o parâmetro θ é uma função suave do vector das médias.

¹ Peter Hall(1988) p.943

A amplitude do intervalo $[\hat{\theta}_{bc}(\alpha), \hat{\theta}_{bc}(1-\alpha)]$ tem como expressão,

$$\begin{aligned} A_{bc}(1-2\alpha) &= 2n^{-1/2} \hat{\sigma} \{ z_{1-\alpha} + n^{-1} [\hat{p}_{21}(z_{1-\alpha}) - 2\hat{p}'_1(z_{1-\alpha}) \hat{p}_1(0)] + O_p(n^{-5/2}) \} \\ &= 2n^{-1/2} \hat{\sigma} z_{1-\alpha} + O_p(n^{-3/2}) \end{aligned} \quad (3.36)$$

Nesta fórmula é patente que a correcção do enviesamento tem um efeito directo (através de $\hat{p}_1(0)$) na amplitude do intervalo. No entanto, a amplitude do intervalo BC difere da amplitude do intervalo correcto por um termo $O_p(n^{-3/2})$, verificando-se assim um erro de grandeza idêntica à do erro de amplitude dos I.C. *standard* e do método dos percentis. É de notar que sendo $\hat{p}_1(0) = 0$ os intervalos BC e método dos percentis têm igual amplitude.

Os I.C. BC podem apresentar ainda uma certa falta de precisão em relação aos I.C. exactos pelo facto de que em muitas situações não se verificam os pressupostos assumidos em (3.30). Na verdade, mesmo existindo a transformação normalizadora g que torna $g(\hat{\theta}) - g(\theta)$ uma variável pivô, não é fácil que $g(\hat{\theta}^*) - g(\hat{\theta})$ tenha distribuição idêntica a $g(\hat{\theta}) - g(\theta)$, sendo antes na maior parte dos casos uma distribuição aproximada. A qualidade da aproximação entre as duas distribuições depende da estrutura da reamostragem. Nomeadamente, quando a reamostragem é feito a partir da distribuição empírica \hat{F} depende da forma como \hat{F} aproxima F_θ (com $\hat{\theta}$ estimativa da M.V.), forma que particularmente em pequenas amostras denota uma certa precariedade.

Uma violação de maior gravidade das hipóteses estabelecidas refere-se aos casos em que não é possível determinar g normalizadora de forma que $g(\hat{\theta}) - g(\theta)$ seja uma variável pivô, isto é, tenha a mesma distribuição Normal quaisquer que sejam os valores de θ . Shenker(1985) introduz este problema através da determinação de I.C. para $\theta = \sigma^2$ em população com distribuição Normal. No estudo de simulação que elaborou conclui que a não

existência de uma transformação normalizadora que determine uma variável pivô afecta nitidamente a *performance* dos I.C. *bootstrap* (método dos percentis e BC).

Este problema prende-se com a oposição que existe, neste tipo de transformações, entre estabilização da variância e normalização. Em termos gerais, o problema caracteriza-se por, dada uma família de distribuições F_θ , encontrar uma transformação $g(x)$ de forma que $g(X)$ tenha distribuição Normal qualquer que seja o valor do parâmetro θ . verificando deste modo,

$$F_\theta(x) = \Phi\left(\frac{g(x) - v_\theta}{\tau_\theta}\right),$$

com v_θ a mediana de $g(X)$ e τ_θ o seu desvio padrão. Acontece frequentemente que o desvio padrão de $g(X)$ não é constante dependendo do valor particular de θ . Como exemplo refere-se o caso da família Poisson contínua analisado em Efron(1982), onde a transformação normalizadora não estabiliza a variância, ao contrário da transformação \tanh^{-1} normalizadora do coeficiente de correlação que tem um desvio padrão constante (independente do parâmetro em questão).

3.3 - MÉTODO COM CORRECÇÃO DO ENVIESAMENTO E DA ACELERAÇÃO

A correcção do enviesamento e da aceleração determina que o ponto crítico, com base no qual se definem os extremos do I.C. para θ , seja definido a partir do quantil de ordem β_α da distribuição *bootstrap* de $\hat{\theta}$. com,

$$\beta_\alpha = \Phi \left[z_0 + \frac{z_0 + z_\alpha}{1 - \alpha(z_0 + z_\alpha)} \right] , \quad (3.37)$$

onde z_0 e a são respectivamente as constantes de enviesamento e aceleração, que serão definidas posteriormente. Finalmente, o ponto crítico mencionado vem,

$$\hat{\theta}_{bca}(\alpha) = \hat{\theta}_p(\beta_\alpha) = \hat{G}^{-1} \left[\Phi \left(z_0 + \frac{z_0 + z_\alpha}{1 - \alpha(z_0 + z_\alpha)} \right) \right] . \quad (3.38)$$

O I.C. para θ unilateral a $\alpha \times 100\%$ BCa é igual a $[-\infty, \hat{\theta}_{bca}(\alpha)]$, enquanto que o intervalo bilateral a $(1-2\alpha) \times 100\%$ corresponde a $[\hat{\theta}_{bca}(\alpha), \hat{\theta}_{bca}(1-\alpha)]$.

O grande problema deste método consiste na determinação da constante a , constante de aceleração. A dedução de a depende das diversas abordagens teóricas que validam o método BCa, embora, conforme se irá concluir, em determinadas circunstâncias as várias definições de a sejam equivalentes.

3.3.1 - Abordagem com base numa transformação normalizadora

A) A formalização de Efron

Hipóteses: Como já foi referido, o método BC assume a hipótese restritiva de que a transformação normalizadora estabiliza simultaneamente a variância quando, em grande parte das situações, se verifica uma contradição entre normalização e estabilização da variância - a transformação que permite uma maior normalização (uma distribuição mais próxima da Normal) produz um desvio padrão que varia¹ com θ .

Efron dá mais alcance aos I.C. BCa quando introduz uma transformação g monótona que, para,

$$\phi = g(\theta) \quad , \quad \hat{\phi} = g(\hat{\theta}) \quad \text{e} \quad \hat{\phi}^* = g(\hat{\theta}^*) \quad , \quad (3.39a)$$

verifica,

$$\frac{\hat{\phi} - \phi}{\tau_{\hat{\phi}}} \sim \mathcal{N}(-z_0, 1) \quad \text{e} \quad \frac{\hat{\phi}^* - \hat{\phi}}{\tau_{\hat{\phi}}} \sim \mathcal{N}(-z_0, 1) \quad , \quad (3.39)$$

sendo $\tau_{\hat{\phi}} = 1 + \alpha \phi$, α e z_0 determinadas constantes.

Estas hipóteses são mais latas que as assumidas no BC dado que o desvio padrão de $g(\hat{\theta})$ deixa de ser estável, passando a depender do valor particular assumido pelo parâmetro desconhecido θ através de $g(\theta)$.

A constante z_0 representa a correcção do enviesamento e é definida por (3.26)² enquanto que α é a chamada constante de aceleração. A interpretação desta constante merece ser mencionada constituindo objecto das próximas linhas.

¹ Em contrapartida, a transformação que produz um desvio padrão estável não permite uma normalização nítida da variável. Sobre este assunto veja-se Efron(1982b), nomeadamente o exemplo referente à família Poisson contínua.

² Dado que, neste caso, z_0 segue um raciocínio dedutivo igual ao verificado no método B.C.

Interpretação de a : Para compreender o significado de a suponha-se que o parâmetro desconhecido assume um dado valor θ_0 , conduzindo ao valor ϕ_0 da função $\phi = g(\theta)$. Quando se verifica uma modificação de ϕ a partir de ϕ_0 , o desvio padrão de $\hat{\phi}$ varia na quantidade $\tau_{\hat{\phi}} - \tau_{\hat{\phi}_0} = a(\phi - \phi_0)$, donde,

$$d\left(\frac{\tau_{\hat{\phi}}}{\tau_{\hat{\phi}_0}}\right) = a \times d\left[\frac{(\phi - \phi_0)}{\tau_{\hat{\phi}_0}}\right],$$

e, desta forma,

$$a = \frac{d(\tau_{\hat{\phi}}/\tau_{\hat{\phi}_0})}{d[(\phi - \phi_0)/\tau_{\hat{\phi}_0}]}.$$

De acordo com esta expressão, a representa a variação relativa do desvio padrão de $g(\hat{\theta})$ em relação à variação de ϕ por unidade de desvio padrão. Mais complexa é a determinação da constante de aceleração, assunto abordado nos próximos parágrafos.

Determinação de a : Na determinação de a há que distinguir duas situações que assumem especificidades próprias - os problemas paramétricos e os problemas não paramétricos.

Começando por estudar os problemas paramétricos, vai supôr-se numa primeira abordagem de forma a simplificar os resultados, que a família de distribuições só apresenta um parâmetro desconhecido, θ . (ou seja, é uma família onde o espaço dos parâmetros é unidimensional). Posteriormente os resultados serão generalizados para famílias caracterizadas por parâmetro k -dimensional ou vector de k parâmetros, $\underline{\eta}$, desconhecido em que $\theta = t(\underline{\eta})$, verificando-se $\hat{\theta} = t(\hat{\underline{\eta}})$.

¹ t é uma função de \mathbb{R}^k em \mathbb{R} .

Nas famílias de parâmetro unidimensional a determinação de α é feita a partir da função *score* da distribuição de $\hat{\phi}$. $i_{\phi}(\hat{\phi})$,¹ definida por,

$$i_{\phi}(\hat{\phi}) = \frac{\partial}{\partial \phi} \log f_{\phi}(\hat{\phi}) .$$

Com base em (3.39) pode escrever-se,

$$\hat{\phi} = \phi + \tau_{\phi}(Z - z_0) \quad , \quad Z \sim \mathcal{N}(0, 1) .$$

Sabendo ainda que $\tau_{\phi} = 1 + \alpha \phi$, Efron conclui que $i_{\phi}(\hat{\phi})$ satisfaz,

$$\tau_{\phi} i_{\phi}(\hat{\phi}) \sim (1 - \alpha z_0) \left[Z + \frac{\alpha}{1 - \alpha z_0} (Z^2 - 1) \right] \quad (3.40)$$

Nesta situação, a partir dos momentos de Z , (3.40) verifica,

$$\frac{1}{6} \text{skew}(i_{\phi}) = \epsilon_0 \frac{1 + (4/3)\epsilon_0^2}{(1 + 2\epsilon_0^2)^{3/2}} , \quad (3.41a)$$

sendo $\epsilon_0 = \alpha/(1 - \alpha z_0)$ e $\text{skew}(X)$ representa o coeficiente de assimetria da v.a. X (igual a $\mu_3(X)/\mu_2(X)^{3/2}$). Efron(1987) mostra que α e z_0 são ambos $O(n^{-1/2})$ possibilitando que,

$$\epsilon_0 = \alpha [1 + O(n^{-1})] . \quad (3.41b)$$

Dado que $\text{skew}(i_{\phi}) = \text{skew}(i_{\theta})$,² onde i_{θ} representa a função *score* da família de distribuições de $\hat{\theta}$. é fácil de verificar a partir de (3.41a) e (3.41b),

$$\frac{1}{6} \text{skew}_{\theta}(i_{\theta}) \doteq \alpha [1 + O(n^{-1})] .$$

Esta última expressão sugere uma boa aproximação para o cálculo de α através de,

$$\alpha \approx \frac{1}{6} \text{skew}_{\theta=\hat{\theta}} \left\{ \frac{\partial [\log f_{\theta}(\hat{\theta})]}{\partial \theta} \right\} \quad (3.42)$$

¹ Veja-se Efron(1987) p.175

² Esta igualdade verifica-se porque ϕ resulta de uma transformação monótona sobre θ , donde $i_{\phi}(\hat{\phi})$ e $i_{\theta}(\hat{\theta})$ são proporcionais.

A substituição em (3.42) de θ desconhecido por $\hat{\theta}$, estimativa da M.V., causa apenas erros de terceira ordem não afectando assim as propriedades de segunda ordem dos I.C. BCa.

É de notar que em muitos casos torna-se conveniente determinar a constante a substituindo em (3.42) i_0 por $\partial L_0(\underline{y})/\partial \theta$ com $\underline{y} = (x_1, x_2, \dots, x_n)$ e $\partial L_0(\underline{y})/\partial \theta$ a função *score* dos dados \underline{y} , correspondendo $L_0(\underline{y})$ ao logaritmo da função de verosimilhança¹.

Para determinar a constante de aceleração quando a família de distribuições é caracterizada por um vector de parâmetros desconhecidos, $\underline{\eta}$ há que recorrer ao artifício de transformar esta família multiparâmetros numa família com apenas um parâmetro desconhecido. Desta forma, torna-se possível aplicar ao problema os resultados atrás deduzidos.

A família $F_{\underline{\eta}}$ será substituída pela subfamília menos favorável de Stein, segundo o procedimento a descrever de seguida. Seja $\hat{\underline{\eta}}$ a estimativa da M.V. de $\underline{\eta}$ e defina-se a matriz $k \times k$ $I_{\underline{\eta}}$ e o vector $k \times 1$ $\underline{\nabla}$ como

$$I_{\underline{\eta}} = \left[\frac{\partial^2 \log f_{\underline{\eta}}}{\partial \eta_i \partial \eta_j} \bigg|_{\underline{\eta} = \hat{\underline{\eta}}} \right]_{i,j} \quad \text{e} \quad \underline{\nabla} = \left[\frac{\partial t(\underline{\eta})}{\partial \eta_i} \bigg|_{\underline{\eta} = \hat{\underline{\eta}}} \right]_i.$$

A direcção menos favorável que passa por $\hat{\underline{\eta}} = \hat{\underline{\eta}}$ é dada por $\underline{\hat{u}} = I_{\underline{\eta}}^{-1} \underline{\nabla}$. A família menos favorável F^* é a subfamília de $F_{\underline{\eta}}$ unidimensional que passa por $\underline{\eta}$ com direcção $\underline{\hat{u}}$ a que corresponde a densidade $f_{\underline{\eta} + \lambda \underline{\hat{u}}}$. Esta resulta assim de substituir o vector de parâmetros $\underline{\eta}$ desconhecido pelo vector $\hat{\underline{\eta}} + \lambda \underline{\hat{u}}$ com $\lambda \in \mathbb{R}$. Dado que $\hat{\underline{\eta}}$ e $\underline{\hat{u}}$ são conhecidos, λ passa a ser o parâmetro unidimensional desconhecido da família de distribuições. Esta família é a menos favorável porque a informação para $\theta(\lambda) = t(\hat{\underline{\eta}} + \lambda \underline{\hat{u}})$ em $\lambda = 0$ é idêntica à informação de $\theta = t(\underline{\eta})$ no problema original. Qualquer subfamília de $F_{\underline{\eta}}$ definida a partir de uma direcção diferente de $\hat{\underline{\eta}} + \lambda \underline{\hat{u}}$ tem

¹ Veja-se Efron(1987) p.184

uma informação de Fisher superior para θ . Assim escolhe-se a subfamília menos favorável para que a estimação de θ não seja artificialmente facilitada.

Ao problema transformado de um só parâmetro é possível aplicar a expressão (3.42) que determina $i(\theta)$ pela função score dos dados). Deste modo, uma boa aproximação para a constante de aceleração em problemas multiparamétricos é dada por,

$$\alpha = \frac{1}{6} skew_{\lambda=0} \left[\frac{\partial L_{\hat{\eta} + \lambda \underline{u}}}{\partial \lambda} \right], \quad (3.43)$$

em que $L_{\hat{\eta} + \lambda \underline{u}}$ corresponde ao logaritmo da função de verosimilhança em $\underline{y} = (x_1, x_2, \dots, x_n)$ para a família de densidades $f_{\hat{\eta} + \lambda \underline{u}}$.

A fórmula (3.43) assume uma expressão particularmente simples para a família exponencial, expressão que é de todo o interesse desenvolver dada a relevância desta família.

Uma f.d.p. pertencente à família exponencial pode ser caracterizada por,

$$f_{\underline{\eta}}(\underline{y}) = \exp \{ [\underline{\eta}' \underline{y} - \Psi(\underline{\eta})] \} f_0(\underline{y}).$$

Passando para a respectiva subfamília menos favorável substituindo $\underline{\eta}$ por $\hat{\underline{\eta}} + \lambda \underline{u}$ e aplicando a fórmula (3.43), obtém-se a expressão,¹

$$\alpha = \frac{1}{6\sqrt{n}} \frac{\Psi^{(3)}(0)}{(\Psi^{(2)}(0))^{3/2}}, \quad (3.44)$$

$$\text{com } \Psi^{(j)}(0) = \left. \frac{\partial^j \Psi(\hat{\underline{\eta}} + \lambda \underline{u})}{\partial \lambda^j} \right|_{\lambda=0}.$$

¹ Veja-se Efron(1987) p.177.

Nos problemas não paramétricos a família F_{η} é totalmente desconhecida. Esta ignorância torna o vector η de dimensão infinita (dado que F_{η} engloba todas as distribuições possíveis) conduzindo a um problema insolúvel. Para ultrapassar este impasse é necessário impôr que F_{η} tenha suporte apenas em \hat{X} , conjunto dos valores observados.

Nestas circunstâncias, F_{η} resume-se a uma família de distribuições com n parâmetros,

$$\eta_1, \eta_2, \dots, \eta_n \quad \text{com} \quad \eta_i = P(X=x_i), \quad \eta_i \geq 0 \quad \text{e} \quad \sum_{i=1}^n \eta_i = 1.$$

Ao vector $\eta^0 = (1/n)\underline{e} = [1/n \dots 1/n]$ corresponde $F_{\eta^0} = \hat{F}$ a distribuição empírica da amostra.

Com η desconhecido, a inferência para $\theta = t(\eta)$ tem de ser processada com base numa realização hipotética do universo fixo $\hat{X} = (x_1, x_2, \dots, x_n)$. Seja $\underline{y}^* = (x_1^*, x_2^*, \dots, x_n^*)$. Esta realização é caracterizada probabilisticamente pelo vector $\underline{f}^* = [f_1^*, f_2^*, \dots, f_n^*]$ $f_i^* = \# \{X_j^* = x_i\} / n$. Nestas circunstâncias o vector \underline{f}^* tem distribuição multinomial, $\underline{f}^* \sim \text{Mult}_n(n, \underline{\eta}) / n$ enquanto que $F_{\hat{\eta}} \sim \text{Mult}_n(n, \underline{\eta}^0) / n$.

Em termos gerais, impondo a restrição mencionada sobre o universo do problema, este passa a caracterizar-se por uma família de distribuições F_{η} do tipo, $F_{\eta} \sim \text{Mult}_n(n, \underline{\eta}) / n$. Desta forma, o problema não paramétrico é transformado num problema paramétrico vulgar, para o qual é já conhecida a expressão que determina a constante de aceleração. Assim, a dedução da constante de aceleração em problemas não paramétricos é muito simplesmente obtida concretizando as fórmulas conhecidas para a família de distribuições $\text{Mult}_n(n, \underline{\eta}) / n$.

A família de distribuições $Mult_n(n, \underline{\eta})/n$ faz parte da família exponencial. Como tal, a subfamília de $F_{\underline{\eta}}$ menos favorável que passa em $\hat{\underline{\eta}}$ corresponde a

$$F_{\lambda} \sim mult_n(n, \underline{W}^{\lambda})/n \quad \text{com} \quad W_i^{\lambda} = \frac{e^{\lambda U_i}}{\sum_{j=1}^n e^{\lambda U_j}}$$

sendo U_i a função de influência empírica de $\hat{\theta} = \theta(F)$, que se define por

$$U_i = \lim_{\epsilon \rightarrow 0} \frac{\theta[(1-\epsilon)\hat{F} + \epsilon \delta_i] - \theta(\hat{F})}{\epsilon}, \quad (3.45)$$

em que δ_i representa uma distribuição degenerada que atribui massa 1 a x_i . Por sua vez, a dedução de a é processada muito simplesmente coma aplicação a F_{λ} da fórmula (3.44).

Como resultado, Efron apresenta uma boa aproximação para a expressa por,

$$a \approx \frac{1}{6} \left[\frac{\sum_{i=1}^n U_i^3}{\left(\sum_{i=1}^n U_i^2 \right)^{3/2}} \right] \quad (3.46)$$

São raros os casos em que U_i é calculado de forma exacta. Regra geral, determina-se U_i atribuindo a ϵ um valor pequeno. Usualmente este valor é $-1/(n-1)$. Com $\epsilon = -1/(n-1)$ U_i vem,

$$U_i = (n-1)[\hat{\theta}(\cdot) - \hat{\theta}(i)] \quad , \quad \hat{\theta}(\cdot) = \frac{1}{n} \sum_{i=1}^n \hat{\theta}(i) \quad , \quad (3.47)$$

sendo $\hat{\theta}(i)$ a estimativa obtida (através do estimador $\hat{\theta}$) excluindo a observação i da amostra inicial.

1 Sobre este assunto veja-se Efron(1987) secções 7 e 8.

Se o universo for na realidade finito, $X = \{1, \dots, L\}$ então os I.C. BCa não paramétricos têm propriedades iguais às dos paramétricos.

Relação entre z_0 e a : As constantes de enviesamento e aceleração, respectivamente z_0 e a , verificam uma relação importante em problemas de um só parâmetro. Efron(1987) p.182, demonstra que sendo $\hat{\theta}$ estimador da M.V. de θ com bias $O(n^{-1})$, desvio padrão $O(n^{-1/2})$, coeficiente de assimetria $O(n^{-1/2})$ e coeficiente de Kurtosis $O(n^{-1})$, então verifica-se

$$z_0 \approx a \quad (3.48)$$

Dedução do I.C.: A dedução do I.C. para θ tendo subjacentes os pressupostos (3.39a) e (3.39) recorre a um raciocínio idêntico ao aplicado no método dos percentis e no BC; só que neste caso assume uma maior complexidade dado que a variância de $\hat{\phi}$ deixa de ser constante passando a depender do próprio parâmetro θ . Este raciocínio resume-se nos seguintes passos: inicialmente determina-se o I.C. para ϕ utilizando a variável pivô Normal $(\hat{\phi} - \phi)/\tau$, e depois transformam-se os extremos deste intervalo através de g^{-1} nos extremos do I.C. para θ .¹ Com este procedimento obtêm-se os intervalos, unilateral a $\alpha \times 100\%$. $[-\infty, \hat{\theta}_{bc\alpha}(\alpha)]$, e bilateral a $(1-2\alpha) \times 100\%$. $[\hat{\theta}_{bc}(\alpha), \hat{\theta}_{bc}(1-\alpha)]$, com $\hat{\theta}_{bc\alpha}(\alpha)$ dado por (3.38). Verificando-se as condições (3.39), estes intervalos são exactos. No entanto, em geral estas condições verificam-se apenas aproximadamente. Por outro lado, as deduções teóricas da constante de aceleração apresentadas por Efron são apenas aproximações convenientes da verdadeira constante a . Perante estes factos, a determinação de intervalos BCa exactos fica bastante comprometida, o que não exclui contudo o facto destes intervalos terem propriedades assintóticas assinaláveis, conforme será analisado posteriormente.

¹ Veja-se Efron(1987) p.174.

B) A variante de Diciccio e Tibshirani

A variante de BCa, desenvolvida por Diciccio e Tibshirani, o denominado BCa^0 , está imbuída de perspectiva diferente da assumida por Efron.

Os I.C. de Efron são determinados a partir da distribuição *bootstrap* de $\hat{\theta}, \hat{G}(s)$, seleccionando-se os percentis de acordo com o nível nominal de confiança combinado com as correcções do enviesamento e aceleração; postula a existência de uma transformação normalizadora g que contudo não necessita de ser conhecida. Pelo contrário, na variante BCa^0 a transformação g assume um papel fulcral na determinação dos extremos do I.C., escamoteando-se o conhecimento da distribuição *bootstrap*.

A lógica do método compreende os seguintes passos:

- 1) Determinar a transformação monótona $g_1(t)$ que estabiliza a variância assymptoticamente;
- 2) Deduzir a transformação $g_a(s)$ reductora da assimetria;
- 3) Combinar $g_1(t)$ e $g_a(s)$ originando a transformação monótona normalizadora, $g(t) = g_a(g_1(t))$ que verifica aproximadamente as condições expressas em (3.39);
- 4) Conhecendo a expressão analítica de $g(t)$ é possível determinar um I.C. para θ - depois de deduzido o I.C. para $g(\theta)$ (recorrendo à estrutura probabilística expressa em (3.39)), obtém-se o I.C. para θ invertendo os extremos do primeiro através de g^{-1} .

Como o BCa^0 assenta nos mesmos pressupostos que o BCa, as constantes z_0 e a são definidas da mesma forma que em Efron.

Para determinar o I.C. BCa^0 não é necessário o conhecimento da distribuição *bootstrap* de $\hat{\theta}$. $\hat{G}(s)$, desde que se calcule z_0 de acordo com uma fórmula aproximada que envolve as funções empíricas de influência de primeira e segunda ordem. Esta metodologia tem a vantagem de recorrer apenas a $n+2$ cálculos adicionais da estatística (número que é substancialmente inferior ao subjacente à metodologia de Efron).

Construção de g : Abordam-se de início os problemas em que o espaço do parâmetro é unidimensional, generalizando-se de seguida para espaços multidimensionais, e por fim, analisa-se a situação não paramétrica.

Seja $L_0(\underline{y})$ o logaritmo da verosimilhança para θ com base na amostra \underline{y} , e $\hat{\theta}$ o estimador da M.V. para θ . Seja ainda $k_2(\theta) = -E\{d^2 L_0(\underline{y})/d\theta^2\}$ a esperança da Informação de Fisher para θ .

Nestas condições, a transformação $g_1(t)$ é definida através da expressão,

$$g_1(t) = \int_c^t k_2(u)^{1/2} du,$$

com c uma constante arbitrária.

Por sua vez, a transformação $g_a(s)$ é dada por,

$$g_a(s) = \frac{e^{as} - 1}{a},$$

onde a representa a constante de aceleração.

Por fim, a transformação normalizadora $g(t)$ é igual a,

$$g(t) = g_a(g_1(t)) = \frac{\exp\left(\alpha \int_c^t k_2(u)^{1/2} du\right) - 1}{\alpha},$$

e verifica,¹

$$E\{g(\hat{\theta}) - g(\theta)\} = -z_0[1 + \alpha g(\theta)] + O(n^{-1})$$

$$Var\{g(\hat{\theta}) - g(\theta)\} = [1 + \alpha g(\theta)]^2 + O(n^{-1}),$$

de forma a ser coerente com as hipóteses expressas em (3.39) com um erro de grandeza $1/n$. Na construção de $g(t)$ é necessário o conhecimento da constante de aceleração, que será determinada de acordo com (3.42).

É de notar que, na determinação de $g(t)$ a variância é desestabilizada, mas a variância resultante, $[1 + \alpha g(\theta)]^2 = \tau_g^2$, permite ainda que seja aplicada uma abordagem com base na variável pivô $(g(\hat{\theta}) - g(\theta))/\tau_g$.

O extremo de ordem α do I.C. $BC\alpha^0$, $\hat{\theta}_{bca^0}(\alpha)$ é dado por,

$$\hat{\theta}_{bca^0}(\alpha) = g^{-1}\left[\frac{z_0 + z_\alpha}{1 - \alpha(z_0 + z_\alpha)}\right], \quad (3.49)$$

definindo o I.C. unilateral a $\alpha \times 100\%$, como $[-\infty, \hat{\theta}_{bca^0}(\alpha)]$, e o I.C. bilateral a $(1 - 2\alpha) \times 100\%$, como $[\hat{\theta}_{bca^0}(\alpha), \hat{\theta}_{bca^0}(1 - \alpha)]$.

Nos problemas multiparâmetros a dedução de $g(t)$ tem subjacente o artifício usado por Efron, consistindo em reduzir a família F_3 à subfamília, de um só parâmetro, menos favorável de Stein. A esta subfamília corresponde a densidade $f_{\hat{\eta} - \lambda \underline{u}}$ com $\hat{\eta}$ estimativa da M.V. e \underline{u} definido como anteriormente.

¹ Diccio e Tibshirani(1987) p.166

Assim, a transformação $g_1(t)$ define-se por,

$$g_1(t) = \int_c^t [k_2^\lambda(u)]^{1/2} du ,$$

sendo $k_2^\lambda(u)$ a esperança de Fisher para λ na família $f_{\hat{\eta} + \lambda \hat{\underline{u}}}$. A transformação $g_a(s)$ mantém a expressão relativa ao caso anterior,

$$g_a(s) = \frac{e^{as} - 1}{a} .$$

Enquanto que a transformação $g(t)$ toma a forma,

$$g(t) = g_a(g_1(t)) = \frac{\exp\left(\alpha \int_c^t [k_2^\lambda(u)]^{1/2} du\right) - 1}{\alpha} ,$$

verificando ainda as condições (3.48).

Depois de conhecido $g(t)$, e consequentemente $g^{-1}(t)$; calculando a de acordo com (3.43) e z_0 segundo o método que irá ser apresentado ou segundo a forma clássica de Efron, é possível determinar um I.C. para λ . Este intervalo facilmente se transforma num I.C. para θ através da relação $\theta(\lambda) = t(\hat{\eta} + \lambda \hat{\underline{u}})$.

Nos problemas não paramétricos, utilizam-se os artifícios que Efron referiu no BCa para definir uma família fictícia de distribuições de um só parâmetro. Esta família corresponde a

$$F_\lambda \sim mult_n(n, W^\lambda) / n \quad \text{com} \quad W_i^\lambda = e^{\lambda U_i} / \sum_{j=1}^n e^{\lambda U_j} ,$$

com U_i a função de influência empírica definida em (3.45).

Para esta família a esperança da Informação de Fisher assume a forma:¹

$$k_2^\lambda(s) = n \left[\frac{\sum_{i=1}^n U_i^2 e^{U_i s}}{\sum_{i=1}^n e^{U_i s}} - \left(\frac{\sum_{i=1}^n U_i e^{U_i s}}{\sum_{i=1}^n e^{U_i s}} \right)^2 \right] .$$

¹ Dicciccio e Tibshirani(1987) p.168

Com $k_2^h(s)$ assim definido, $g_1(t)$ calcula-se facilmente utilizando uma regra de integração numérica. Note-se que $k_2^h(s)$ é uma função não negativa permitindo que $g_1(t)$ seja monótona crescente e invertível. A função $g_a(s)$ mantém-se como anteriormente, sendo neste caso determinado a a partir da expressão (3.46) que Efron definiu para situações não paramétricas.

A função $g_1(t)$ é, geralmente, difícil de calcular. No entanto, tal como a constante de aceleração, apresenta uma expressão muito mais simples e operacional nos problemas não paramétricos (pois $k_2^h(s)$ só depende das funções empíricas de influência, não sendo afectado pela complexidade da estatística).

Relação entre $\hat{\theta}_{bca^0}$ e $\hat{\theta}_{bca}$: Diccio e Tibshirani(1987) mostram que,

$$\frac{\hat{\theta}_{bca^0}(\alpha) - \hat{\theta}_{bca}(\alpha)}{\hat{\sigma}} = O_p(n^{-1}) ,$$

$$\text{com } \hat{\sigma} = [k_2(\hat{\theta})]^{-1/2} .$$

Este resultado denota que o $BC\alpha^0$ conserva as propriedades de segunda ordem do BCa.

Determinação de z_0 com $n+2$ iterações: Seja $\underline{U} = [U_1 \dots U_t \dots U_n]$ o vector que contém as funções empíricas de influência das observações, com U_i definido de acordo com (3.45); e seja V a matriz $n \times n$ das funções empíricas de influência de segunda ordem. Uma boa aproximação para a correcção do enviesamento, z_0 , é dada por,

$$z_0 \approx \Phi^{-1} [2\Phi(z_{01})\Phi(z_{02})] ,$$

$$\text{em que } z_{01} = \frac{1}{6} \times \frac{\sum_{i=1}^n U_i^3}{\left(\sum_{i=1}^n U_i^2\right)^{3/2}}$$

$$\text{e } z_{02} = [U^t V U / |U|^2 - \text{tr}(V)] / (2n |U|^2) ,$$

(z_{01} é igual à aproximação para a na situação não paramétrica).

Para calcular z_0 não é necessário o conhecimento da matriz V , dado que existe uma expressão operacional para o cômputo de $U^t V U$ e de $\text{tr} V$ sem envolver as influências de segunda ordem. Denomine-se $U(i, \epsilon)$ a função empírica de influência (3.45) para um dado ϵ escolhido. Fazendo apelo à notação já utilizada no capítulo 1 deste trabalho, considere-se $\hat{\theta}(\underline{P}^*)$ como representando o valor da estatística para a amostra resultante do vector de reamostragem \underline{P}^* que atribui massa P_i^* a x_i .

As quantidades $U^t V U$ e $\text{tr} V$ são obtidas de forma aproximada através de¹

$$U^t V U \approx [\hat{\theta}(\underline{P}^0 + \epsilon U) + \hat{\theta}(\underline{P}^0 - \epsilon U) - 2\hat{\theta}(\underline{P}^0)] / \epsilon^2 ,$$

$$\text{Tr} V \approx \frac{1}{\epsilon^2} \sum_{i=1}^n (U_i, \epsilon) ,$$

em que \underline{P}^0 é o vector de reamostragem que corresponde a \hat{F} , função de distribuição empírica da amostra, donde $\hat{\theta}(\underline{P}^0) = \hat{\theta}$.

Em resumo, as $n+2$ computações necessárias a este procedimento são, fixado um dado ϵ , n relativas às funções empíricas de Influência, $U(1, \epsilon), \dots, U(n, \epsilon)$, e as outras duas referem-se a $\hat{\theta}(\underline{P}^0 + \epsilon U)$ e $\hat{\theta}(\underline{P}^0 - \epsilon U)$.

¹ Diccio e Tibshirani(1987) p.165

Determinação de uma aproximação a $\hat{G}(s)$: A metodologia que determina os I.C. BCa^0 permite simultaneamente deduzir uma aproximação da distribuição *bootstrap* para $\hat{\theta}$, $\hat{G}(s)$, sem recorrer às reamostragens *bootstrap*.

Dado que ,

$$\hat{G}^{-1} \left[\Phi \left(z_0 + \frac{z_0 + z_a}{1 - \alpha(z_0 + z_a)} \right) \right] \quad \text{e} \quad g^{-1} \left[\frac{z_0 + z_a}{1 - \alpha(z_0 + z_a)} \right] ,$$

diferem apenas por $O_p(n^{-1})$ é possível determinar $\hat{G}^{-1}(p)$ a partir de g^{-1} conhecido, qualquer que seja o p que se considere. O procedimento estrutura-se de acordo com o seguintes passos:

- 1) Para cada p determinar z_a tal que,

$$p = \Phi \left(z_0 + \frac{z_0 + z_a}{1 - \alpha(z_0 + z_a)} \right) \quad \text{donde} \quad z_a = \frac{\Phi^{-1}(p) - z_0}{1 + \alpha(\Phi^{-1}(p) - z_0)} - z_0 ;$$

- 2) Para o valor z_a encontrado em (1) calcular,

$$g^{-1} \left[\frac{z_0 + z_a}{1 - \alpha(z_0 + z_a)} \right] ,$$

que constitui uma aproximação de $\hat{G}^{-1}(p)$.

Uma densidade, $r(s)$, que aproxima o histograma *bootstrap* será obtida de acordo com¹

$$r(s) = \phi \left[\exp(g_1(s)\alpha) - \frac{1}{\alpha} + z_0 \right] \times \exp(g_1(s)\alpha) [k_2(s)]^{1/2} .$$

Quando o espaço dos parâmetros é multidimensional, $r(s)$ refere-se à densidade de $\hat{\lambda}$. Para obter a densidade de $\hat{\theta}$ basta multiplicar $r(s)$ por $d\hat{\lambda}/d\hat{\theta} = n/k_2^{\lambda}(s)$.

A grande virtude desta abordagem consiste em deduzir a expressão analítica de uma aproximação (de boa qualidade) da distribuição *bootstrap* da estatística, dado que, com a metodologia de Efron, esta distribuição é quase sempre determinada numericamente através de um processo de Monte Carlo.

¹ Diccio e Tibshirani(1987) p.169

Diciccio e Tibshirani apontam algumas vantagens nesta aproximação da função de distribuição *bootstrap*:¹

This approximating procedure can be thought of as a refinement of the usual central limit theorem approximation $N(\hat{\theta}, k_2(\hat{\theta})^{-1})$, correct to order $n^{-1/2}$. The new approximation incorporates three order $n^{-1/2}$ components: $g(\cdot)$, z_0 , and a . In a parametric setting, $(r(s))$ could prove to be a useful alternative to an Edgeworth expansion. It has two distinct advantages over Edgeworth expansions: a) it is always nonnegative because $g(\cdot)$ is monotone increasing and b) is computable (albeit not often by hand) for general first order efficient statistics $\hat{\theta}$.

O BCa^o constitui uma alternativa ao BCa onde não é necessário recorrer às sucessivas reamostragens *bootstrap*. Sobre este método Efron² refere:

Diciccio and Tibshirani BCa^o method requires much less computation, at the expense of a greater reliance on asymptotic approximations.

Apesar de exigir um esforço computacional menor em quantidade os seus cálculos são no entanto mais complexos (pois há que recorrer a integrações, inversão de funções, cálculo da Informação de Fisher, entre outros.)

O grande contributo do BCa^o consiste em mostrar que, para os problemas paramétricos, existe uma transformação monótona normalizadora que verifica as hipóteses postuladas por Efron com um erro de grandeza n^{-1} , não afectando as propriedades de segunda ordem do I.C.. No entanto, para os problemas não

1 Diciccio e Tibshirani(1987) p.170

2 Efron(1987) p.199.

paramétricos os autores recorrem à simplificação de Efron sobre o universo do problema, que constitui uma restrição importante perante os problemas concretos. Persiste assim uma certa vulnerabilidade na fundamentação teórica do BCa em universos não paramétricos.

Como geralmente, em problemas paramétricos, $g_1(t)$ é uma função complicada de determinar, o uso do BCa^o fica praticamente cingido aos problemas não paramétricos (que são no fundo os problemas onde mais se recorre às técnicas *bootstrap*) para os quais as fórmulas fulcrais se obtêm com base nas funções empíricas de influência. A escolha entre BCa e BCa^o depende das opções do utilizador sobre as características dos cálculos a prosseguir.

C) Abordagem de Babu e Konishi

A abordagem de Babu e Konishi distancia-se da abordagem de Efron não tanto a nível dos pressupostos mas, de forma nítida, na determinação das constantes que corrigem do enviesamento e da aceleração.

Tal como Efron, estes autores admitem a existência de uma transformação monótona normalizadora, g , que verifica (3.39), isto é,

$$\frac{g(\hat{\theta}) - g(\theta)}{1 + \alpha g(\theta)} \sim \mathcal{N}(-z_0, 1)$$

No entanto Babu e Konishi tratam um problema mais restricto. Admitem que a estatística $\hat{\theta} = \theta(\hat{F})$ verifica uma expansão estocástica de Von Mises em torno do parâmetro $\theta = \theta(F)$, que se caracteriza da seguinte forma,

$$\theta(F) = \theta(\theta) + \int \theta_1(x, F) dF(x) + \frac{1}{2} \int \int \theta_2(x, y, F) dF(x) dF(y) + o_p(n^{-1})$$

$$\text{com } \int \theta_1(x, F) dF(x) = 0 \quad \text{e} \quad \int \theta_2(x, y, F) dF(y) = 0, \quad \theta_2 \text{ simétrico};$$

sendo θ_i ($i=1,2$) a i -ésima derivada compacta de Hadamard de θ (θ_1 pode ser considerada a função de influência de $\hat{\theta}$).

Os autores consideram que para $\sqrt{n}[\hat{\theta} - \theta]$ a variância, o enviesamento e o coeficiente de assimetria, assintóticos, dependem de θ , sendo respectivamente $\sigma(\theta)$, $\xi(\theta)$, e $\gamma(\theta)$; e admitem ainda que a respectiva distribuição verifica a expansão de Edgeworth

$$P\{\sqrt{n}(\hat{\theta} - \theta) \leq x\sigma(\theta)\} = \Phi(x) - \frac{1}{\sqrt{n}} \left\{ b(\theta) + \frac{1}{6} \gamma(\theta)(x^2 - 1) \right\} \phi(x) + O(n^{-1})$$

$$\text{com } \sigma(\theta)^2 = \text{Var}\{\theta_1(X_1, F)\} = \int \theta_1^2(x, F) dF(x)$$

$$b(\theta) = \frac{1}{2\sigma(\theta)} \int \theta_2(x, x, F) dF(x)$$

$$\gamma(\theta) = \frac{1}{\sigma(\theta)^3} \left\{ \int \theta_1^3(x, F) dF(x) + 3 \int \int \theta_1(x, F) \right.$$

$$\left. \theta_1(y, F) \theta_2(x, y, F) dF(x) dF(y) \right\}.$$

De acordo com estes pressupostos os autores vão construir uma transformação g e estimar a e z_0 de modo que a distribuição de $[g(\hat{\theta}) - g(\theta)]/[1 + ag(\theta)]$ seja aproximadamente Normal estandardizada com um erro de ordem $O(n^{-1})$. Desta forma, a transformação g e as constantes a e z_0 que verificam,

$$P\{g(\hat{\theta}) - g(\theta) \leq (x - z_0)(1 + ag(\theta))\} = \Phi(x) + O(n^{-1}), \quad (3.50)$$

são definidas pelas expressões,

$$g(\hat{\theta}) = \frac{\exp(\alpha \psi(\hat{\theta})) - 1}{\alpha}$$

$$z_0 = -\frac{1}{\sqrt{n}} \left\{ \xi(\theta) - \frac{1}{6} \hat{\gamma}(\theta) \right\}$$

$$\alpha = \frac{1}{\sqrt{n}} \left\{ \hat{\sigma}'(\theta) - \frac{1}{3} \hat{\gamma}(\theta) \right\} \quad (3.51)$$

em que ψ corresponde à solução da equação diferencial $\sigma(\theta)\psi'(\theta) = \sqrt{N}$ e $\sigma'(\theta)$ é a derivada de $\sigma(\theta)$ em ordem a θ .

A expressão apresentada para $g(\hat{\theta})$ é válida se $3\sigma'(\theta) \neq \gamma(\theta)$. Esta expressão é muito semelhante à expressão deduzida por Tibshirani e Diccicio para $g(t)$. Onde está $\psi(\hat{\theta})$ estes autores consideram $\int_0^t [k_2(u)]^{1/2} du$ (com $k_2(u)$ a informação de Fisher esperada) que de alguma forma se relaciona com a expressão dada para ψ visto que $\psi'(\theta) = \sqrt{n}/\sigma(\theta)$. Se $3\sigma'(\theta) = \gamma(\theta)$ g pode ser considerada igual a ψ .

As igualdades que determinam z_0 e α , envolvem as estimativas de $\xi(\theta)$, $\sigma'(\theta)$, e $\gamma(\theta)$, respectivamente $\hat{\xi}(\theta)$, $\hat{\sigma}'(\theta)$, e $\hat{\gamma}(\theta)$, com erros de estimação da ordem $O_p(n^{-1/2})$.

O cálculo de g não é necessário para a determinação do I.C. BCa, dado que o I.C. para θ é obtido directamente a partir da distribuição *bootstrap* para $\hat{\theta}$, envolvendo apenas o conhecimento de α e z_0 . Este I.C. é deduzido a partir do ponto crítico $\hat{\theta}_{bca}(\alpha)$, dado por,

$$\hat{\theta}_{bca}(\alpha) = \hat{G}^{-1} \left[\Phi \left(z_0 + \frac{z_0 + z_\alpha}{1 - \alpha(z_0 + z_\alpha)} \right) \right],$$

que coincide com a expressão do ponto crítico BCa de Efron. Assim, o I.C. unilateral a $\alpha \times 100\%$ corresponde a $[-\infty, \hat{\theta}_{bca}(\alpha)]$ e o I.C. bilateral a $(1-\alpha) \times 100\%$ é dado por $[\hat{\theta}_{bca}(\alpha), \hat{\theta}_{bca}(1-\alpha)]$.

Para os problemas que envolvem um só parâmetro, com $\hat{\theta}$ estimador da M.V. de θ . Babu¹ mostra que se verifica,

$$z_0 = \alpha = \frac{1}{6} \text{skew} \frac{\partial L(\theta)}{\partial \theta}.$$

Neste resultado há que distinguir dois aspectos - para este tipo de problemas, por um lado, a constante de aceleração definida por estes autores é idêntica à constante de aceleração de Efron e por outro lado, existe uma igualdade entre as constantes de aceleração e correcção do enviesamento, sendo coerente com a conclusão (3.48) obtida por Efron.

A partir da verificação de (3.50) conclui-se que os intervalos unilaterais e bilaterais apresentados (com base nas estimativas dadas em (3.51)) têm probabilidade de cobertura real que difere da cobertura nominal por um erro da ordem $O(n^{-1})$. É de notar que, para os I.C. unilaterais este erro é inferior ao verificado com os outros métodos já analisados (percentis e BC).

Da abordagem de Babu e Konishi há que realçar dois aspectos. O primeiro (já referido) reside na virtude de mostrar a possibilidade do BCa produzir I.C. unilaterais com um erro de cobertura inferior ao registado pelos outros I.C. aproximados. O segundo aspecto prende-se com o facto de esta abordagem ser aplicada indiferentemente a universos paramétricos e não paramétricos. Como tal, traduz uma justificação teórica mais consistente da justeza da aplicação do BCa a problemas não paramétricos. No entanto convém recordar que os resultados apresentados têm implícitas as hipóteses de que a estatística verifica uma expansão estocástica de Von Mises e uma expansão de Edgeworth nas condições apresentadas no texto.

1 Babu(1989) p.234 .

3.3.2 - Propriedades do BCa

O ponto crítico $\hat{\theta}_{\text{bca}}(\alpha)$ é uma estimativa *bootstrap* do ponto teórico:

$$\hat{\theta}_{\text{back}}(\beta_a) = \hat{\theta} + n^{-1/2} \hat{\sigma} x_{\beta_a} \quad (3.52)$$

$$\text{com } \beta_a = \Phi\{z_\alpha + 2m + a z_\alpha^2 + O(n^{-1})\}$$

em que $\beta_a = \Phi\{z_\alpha + 2m + a z_\alpha^2 + O(n^{-1})\}$ sendo m definido de acordo com (3.31), e a uma dada constante.

A constante a traduz a correcção adicional ao ponto crítico teórico $\hat{\theta}_{\text{back}}(\beta)$, originando $\hat{\theta}_{\text{back}}(\beta_a)$. Analise-se de seguida a sua dedução. O termo $n^{-1/2}\{p_1(z_\alpha) + q_1(z_\alpha) - 2p_1(0)\}$ da probabilidade de cobertura de $\hat{\theta}_{\text{back}}(\beta)$, expressa em (3.32), pode ser representado por $a z_\alpha^2$, donde resolvendo em ordem a a , vem

$$a = n^{-1/2} z_\alpha^{-2} \{p_1(z_\alpha) + q_1(z_\alpha) - 2p_1(0)\} , \quad (3.53)$$

em que a não depende de z_α .

Esta definição de a induz a que o I.C. unilateral $[-\infty, \hat{\theta}_{\text{back}}(\beta_a)]$ tenha probabilidade de cobertura,¹

$$\pi_{\text{back}}(\beta_a) = P\{\theta \leq \hat{\theta}_{\text{back}}(\beta_a)\} = \alpha + O(n^{-1}) ,$$

reduzindo o erro da probabilidade de cobertura de $O(n^{-1/2})$ para $O(n^{-1})$.

Peter Hall denomina a como constante de aceleração. Este autor apresenta uma outra definição para a , válida pelo menos nos casos mais importantes, e que consiste em considerar a como 1/6 do

¹ Peter Hall(1988) p.935

terceiro momento (assimetria) da aproximação de primeira ordem de $\sqrt{n}(\hat{\theta} - \theta)/\sigma$. Bickel(1988) (discussão contida no texto de Peter Hall(1988)) generaliza heurísticamente este resultado para os casos em que $\hat{\theta} = \theta(\hat{F})$ verifica uma expansão estocástica de Von Mises em torno de $\theta = \theta(F)$, ideia desenvolvida, como foi exposto, por Babu e Konishi.

Peter Hall², concretiza a expressão de a para os problemas em que $\hat{\theta} = f(\bar{X})$ e $\hat{\sigma}^2 = g(\bar{X})$ com \bar{X} multivariada e g e f funções "suaves"³, chegando à conclusão de que⁴

$$\alpha \approx \frac{1}{6} \text{skew} \sqrt{n}(\hat{\theta} - \theta) .$$

O quantil x_{β_a} verifica a expansão,⁴

$$x_{\beta_a} = z_{\alpha} + n^{-1/2} q_1(z_{\alpha}) + O(n^{-1}) ,$$

enquanto que $y_{1-\alpha}$ é igual a,

$$y_{1-\alpha} = z_{1-\alpha} - n^{-1/2} q_1(z_{1-\alpha}) + O(n^{-1}) ,$$

e dado que $z_{1-\alpha} = -z_{\alpha}$ e $q_1(x)$ é função par verifica-se,

$$x_{\beta_a} = -y_{1-\alpha} + O(n^{-1}) .$$

Esta igualdade expressa que a correcção implementada pela constante a (juntamente com a correcção do enviesamento) permite definir um quantil de $H(x)$ que se ajusta ao simétrico do quantil exacto, $-y_{1-\alpha}$, com um erro de ordem n^{-1} , sendo assim correcto

1 Peter Hall(1988) p.935 .

2 Como exemplo de problemas deste tipo refere-se a inferência na função exponencial; estimação não paramétrica da média ou de variâncias ou mesmo do coeficiente de correlação, ou de quocientes ou produtos de médias ou variâncias.

3 Mais concretamente a expressão deduzida pelo autor é equivalente a

$$\alpha = \frac{1[k_3(U) - O(n^{-3/2})]}{6[k_2(U) - O(n^{-1})]^{3/2}} \approx \frac{1}{6} \frac{k_3(U)}{k_2(U)^{3/2}} ,$$

com $k_2(U)$ e $k_3(U)$ respectivamente os segundos e terceiros cumulantes de $\sqrt{n}(\hat{\theta} - \theta)$ sendo $\hat{\theta} = f(\bar{X})$ e $\theta = f(\mu)$.

4 Peter Hall(1988) p.935

de segunda ordem. Desta forma, o extremo teórico aproximado $\hat{\theta}_{back}(\beta_a)$ ajusta-se ao extremo teórico exacto $\hat{\theta}_{stud}(a)$ correctamente até à ordem n^{-1} das respectivas expansões assintóticas (que traduz uma igualdade até ao termo destas correspondente a $1/\sqrt{n}$). Como tal $\hat{\theta}_{back}(\beta_a)$ é correcto de segunda ordem.

Para definir a estimativa *bootstrap* de $\hat{\theta}_{back}(\beta_a)$, $\hat{\theta}_{bca}(a)$, é necessário deduzir a réplica *bootstrap* de β_a e consequentemente de a . A estimativa *bootstrap* da constante de aceleração é dada por,

$$\hat{a} = n^{-1/2} z_a^{-2} \{ \hat{p}_1(z_a) + \hat{q}_1(z_a) - 2\hat{p}_1(0) \} , \quad (3.54a)$$

onde \hat{p}_1 e \hat{q}_1 são os já referidos polinómios que aparecem na expansão de Edgeworth das distribuições *bootstrap* $\hat{H}(x)$ $\hat{K}(x)$, definidas em (3.18). Como já foi apontado, a estimativa *bootstrap* de m corresponde a z_0 (dado por (3.26)), sendo a estimativa *bootstrap* para β_a igual a,

$$\hat{\beta}_a = \Phi \left\{ z_0 + \frac{z_0 + z_a}{1 - \hat{a}(z_0 + z_a)} \right\} .$$

Como, $z_0 + (z_0 + z_a)/[1 - \hat{a}(z_0 + z_a)] = z_a + 2z_0 + \hat{a}z_a^2 + O_p(n^{-1})$, a expressão acima indicada é a réplica directa do β_a teórico definido em (3.52). A definição acima indicada coincide com a definição de Efron para β_a dada em (3.37), diferindo apenas a interpretação dada em cada caso à constante de aceleração.

Hall¹ apresenta a estimativa *bootstrap* da constante de aceleração para os problemas particulares (já referidos) em que $\theta = f(\mu)$, $\hat{\theta} = f(\bar{X})$, $\sigma^2 = g(\mu)$, e $\hat{\sigma}^2 = g(\bar{X})$ com μ e \bar{X} multidi-

¹ Peter Hall(1988) p.939 .

mensionais e f e g funções "suaves". Conclui-se facilmente que neste caso,

$$\hat{\alpha} \approx \frac{1}{6} \text{skew} \sqrt{n}(\hat{\theta}^* - \hat{\theta}) , \quad (3.54)$$

sendo esta definição muito mais operacional do que (3.54a).

Ainda o mesmo autor no texto referido mostra que nos problemas paramétricos de inferência envolvendo a família exponencial (com parâmetro multidimensional), e nos problemas não paramétricos onde o estimador pode ser expresso como função suave de vectores de médias, a dedução para a constante de aceleração que conduz a (3.54) coincide com as expressões apontadas por Efron respectivamente em ambiente paramétrico e em ambiente não paramétrico. Neste contexto, são válidas para os I.C. BCa de Efron as propriedades a apresentar de seguida.

A probabilidade de cobertura do I.C. unilateral a $\alpha \times 100\%$, $[-\infty, \hat{\theta}_{bca}(\alpha)]$ será a estimativa *bootstrap* de $\pi_{back}(\beta_\alpha)$, e é igual a,

$$\pi_{bca}(\alpha) = P\{\theta \leq \hat{\theta}_{bca}(\alpha)\} = \alpha + O_p(n^{-1}) .$$

Esta cobertura é nitidamente melhor do que a cobertura de qualquer um dos I.C. aproximados apresentados (incluindo o *standard*) pois estas eram correctos até à ordem $n^{-1/2}$ enquanto que a primeira é correcta a uma ordem superior, nomeadamente até à ordem n^{-1} .

Nos I.C. bilaterais já não se verifica esta tendência, isto é, a cobertura regista um erro de ordem igual ao verificado nos outros I.C. aproximados analisados. Neste caso, a probabilidade de cobertura do intervalo bilateral a $(1-2\alpha) \times 100\%$ $[\hat{\theta}_{bca}(\alpha), \hat{\theta}_{bca}(1-\alpha)]$ é dada por,

$$\begin{aligned}
\pi_{bca}(1-\alpha) - \pi_{bca}(\alpha) &= (1-2\alpha) - 2n^{-1} \left\{ \frac{1}{2} \hat{q}_1(z_{1-\alpha})^2 z_{1-\alpha} \right. \\
&\quad + \hat{q}_1(z_{1-\alpha}) [q'_1(z_{1-\alpha}) - q_1(z_{1-\alpha}) z_{1-\alpha}] \\
&\quad - q_2(z_{1-\alpha}) - (\hat{p}_1(z_{1-\alpha}) + \hat{q}_1(z_{1-\alpha})) [(\hat{p}_1(z_{1-\alpha}) \\
&\quad + \hat{q}_1(z_{1-\alpha}) - 2\hat{p}_1(0)) z_{1-\alpha}^{-1} - \hat{p}'_1(z_{1-\alpha})] \\
&\quad \left. - \hat{p}_{21}(z_{1-\alpha}) + u z_{1-\alpha} \right\} \phi(z_{1-\alpha}) + O_p(n^{-2}) \\
&= (1-2\alpha) + O_p(n^{-1}) \quad ,
\end{aligned}$$

onde u é uma constante que satisfaz,

$$E \left\{ \left[\sqrt{n} \frac{(\hat{\theta} - \theta)}{\hat{\sigma}} \right] [\sqrt{n} (\hat{q}_1(z_\alpha) - q_1(z_\alpha))] \right\} = u + O(n^{-1}) \quad .$$

Os resultados apresentados sobre o erro da probabilidade de cobertura dos intervalos BCa (pressupondo uma estimativa para a constante de aceleração de acordo com (3.54a)) coincidem com as conclusões obtidas por Babu e Konishi.

A correcção do enviesamento e aceleração não melhora a grandeza do erro patente na amplitude do I.C. aproximado, persistindo em ser $O_p(n^{-3/2})$. Desta forma, amplitude do I.C. BCa bilateral é,

$$\begin{aligned}
A_{bca}(1-2\alpha) &= 2n^{-1/2} \hat{\sigma} \{ z_{1-\alpha} + n^{-1} [\hat{p}_1(z_{1-\alpha}) \\
&\quad + \hat{q}_1(z_{1-\alpha})] [(\hat{p}_1(z_{1-\alpha}) + \hat{q}_1(z_{1-\alpha}) - 2\hat{p}_1(0)) z_{1-\alpha}^{-1} \\
&\quad - \hat{p}'_1(z_{1-\alpha})] + \hat{p}_{21}(z_{1-\alpha}) \} + O_p(n^{-5/2}) \\
&= 2n^{-1/2} \hat{\sigma} z_{1-\alpha} + O_p(n^{-3/2}) \quad .
\end{aligned}$$

Este facto não é inesperado, pois as propriedades de segunda ordem não influem na amplitude do verdadeiro I.C. (dada em (3.10)) visto que esta não depende dos segundos termos da expansão assintótica dos extremos exactos.

O ponto crítico $\hat{\theta}_{bca}(\alpha)$ verifica a expansão,¹

$$\begin{aligned}\hat{\theta}_{bca}(\alpha) = & \hat{\theta} + n^{-1/2} \hat{\sigma} \{ z_\alpha + n^{-1/2} \hat{q}_1(z_\alpha) + n^{-1} \{ (\hat{p}_1(z_\alpha) \\ & + \hat{q}_1(z_\alpha)) [(\hat{p}_1(z_\alpha) + \hat{q}_1(z_\alpha) - 2\hat{p}_1(0)) z_\alpha^{-1} \\ & - \hat{p}'_1(z_\alpha)] + \hat{p}_{21}(z_\alpha) \} \} + O_p(n^{-2}) .\end{aligned}\quad (3.55)$$

Comparando (3.55) com a expansão do extremo exacto $\hat{\theta}_{stud}(\alpha)$ (dada em (3.8)) e sabendo que $\hat{q}_1 = q_1 + O_p(n^{-1/2})$, é possível concluir que,

$$|\hat{\theta}_{bca}(\alpha) - \hat{\theta}_{stud}(\alpha)| = O_p(n^{-3/2}) ,$$

donde se conclui que $\hat{\theta}_{bca}(\alpha)$ é correcto até à segunda ordem.

Efron(1987) estabelece a correcção de segunda ordem do BCa nos problemas de parâmetro unidimensional em que $\hat{\theta}$ tem um comportamento assintótico igual ao de um estimador da M.V., nomeadamente na ordem de grandeza do enviesamento, desvio padrão, assimetria, e *kurtosis*. O resultado de Peter Hall contempla um conjunto mais vasto de situações, situações em que $\hat{H}(x)$ e $\hat{K}(x)$ tenham E.E. e a correcção da aceleração seja fixada de acordo com (3.54a).

¹ Peter Hall(1988) p.943

3.4 - O *BOOTSTRAP-t*

O *Bootstrap-t* consiste em conduzir a inferência a partir da distribuição *bootstrap* da estatística *studentizada* $\sqrt{n}(\hat{\theta} - \theta) / \hat{\sigma}$. Para o desvio padrão de $\hat{\theta}$, $\hat{\sigma}$, deve ser considerada uma estimativa assintoticamente eficiente, de preferência a estimativa da M.V..

Sendo $\hat{K}(x)$ a distribuição *bootstrap* da estatística *studentizada* (conforme a explicitação de (3.18), o I.C. *bootstrap-t* para θ é determinado a partir do ponto crítico,

$$\hat{\theta}_i(\alpha) = \hat{\theta} - \hat{\sigma} n^{-1/2} \hat{y}_{1-\alpha}, \quad (3.56)$$

com $\hat{y}_{1-\alpha}$ o quantil $1-\alpha$ de $\hat{K}(x)$. O I.C. unilateral a $\alpha \times 100\%$ será $(-\infty, \hat{\theta}_i(\alpha))$ enquanto que o I.C. bilateral a $(1-\alpha) \times 100\%$ é igual a $(\hat{\theta}_i(\alpha), \hat{\theta}_i(1-\alpha))$.

É de notar que (3.56) é a estimativa *bootstrap* directa do extremo exacto θ_{stud} . Este método não necessita correcções analíticas, utilizando apenas os percentis de $\hat{K}(x)$.

3.4.1 - Propriedades do *bootstrap-t*

A probabilidade de cobertura do I.C. unilateral a $\alpha \times 100\%$, $[-\infty, \hat{\theta}_i(\alpha)]$ sofre de um erro de grandeza $O(n^{-1})$, tal como a cobertura do I.C. unilateral BCa, sendo igual a,

$$\begin{aligned} \pi_i(\alpha) &= P\{\theta \leq \hat{\theta}_i(\alpha)\} \\ &= \alpha + n^{-1/2} \{ \hat{q}_1(z_\alpha) - q_1(z_\alpha) \} \phi(z_\alpha) + O_p(n^{-1}) \\ &\quad \text{e como } \hat{q}_1(z_\alpha) = q_1(\alpha) + O_p(n^{-1/2}) \quad \text{vem} \\ &= \alpha + O_p(n^{-1}). \end{aligned} \quad (3.57)$$

Por seu turno, o I.C. bilateral a $(1-\alpha) \times 100\%$. $[\hat{\theta}_t(\alpha), \hat{\theta}_t(1-\alpha)]$ tem probabilidade de cobertura dada pela expressão,

$$\begin{aligned} \pi_t(\alpha) - \pi_t(1-\alpha) &= 1 - 2\alpha - 2n^{-1} \left\{ \frac{1}{2} \hat{q}_1(z_{1-\alpha})^2 z_{1-\alpha} + \hat{q}_1(z_{1-\alpha}) [q'_1(z_{1-\alpha}) \right. \\ &\quad \left. - q_1(z_{1-\alpha}) z_{1-\alpha}] - q_2(z_{1-\alpha}) \right. \\ &\quad \left. - \hat{q}_2(z_{1-\alpha}) + u z_{1-\alpha} \right\} \phi(z_{1-\alpha}) + O_p(n^{-2}) \\ &= 1 - 2\alpha + O_p(n^{-1}) \quad , \end{aligned} \quad (3.58)$$

onde u é uma constante que satisfaz,

$$E \left\{ \left[\sqrt{n} \frac{(\hat{\theta} - \theta)}{\hat{\sigma}} \right] [\sqrt{n} (\hat{q}_1(z_\alpha) - q_1(z_\alpha))] \right\} = u + O(n^{-1}) \quad .$$

O *Bootstrap-t* não consegue diminuir a grandeza do erro da probabilidade de cobertura dos I.C. bilaterais, mantendo-se em $O_p(n^{-1})$ tal como nos outros métodos analisados.

A expansão do extremo $\hat{\theta}_t(\alpha)$ conduz à expressão,

$$\hat{\theta}_t(\alpha) = \hat{\theta} + n^{-1/2} \hat{\sigma} \{ z_\alpha + n^{-1/2} \hat{q}_1(z_\alpha) + n^{-1} \hat{q}_{21}(z_\alpha) \} + O_p(n^{-2}) \quad (3.59)$$

Tal como $\hat{\theta}_{bca}(\alpha)$, $\hat{\theta}_t(\alpha)$ é correcto de segunda ordem dado que se verifica,

$$|\hat{\theta}_t(\alpha) - \hat{\theta}_{stud}(\alpha)| = O_p(n^{-3/2}) \quad ,$$

pois $\hat{q}_1 = q_1 + O_p(n^{-1/2})$. No entanto a expansão de $\hat{\theta}_t(\alpha)$ sugere que este ponto tem superioridade em relação a $\hat{\theta}_{bca}(\alpha)$, nas propriedades de terceira ordem dado que é bastante verosímil que \hat{q}_{21} se ajuste mais correctamente a q_{21} do que o polinómio de expressão complexa que caracteriza este termo na expressão de $\hat{\theta}_{bca}(\alpha)$.

A amplitude do I.C. bilateral dado pelo *Bootstrap-t* é igual a,

$$A_i(1-2\alpha) = 2n^{-1/2} \hat{\sigma} \{ z_{1-\alpha} + n^{-1} \hat{q}_{21}(z_{1-\alpha}) \} + O_p(n^{-5/2}) . \quad (3.60)$$

Como $\hat{q}_{21} = q_{21} + O_p(n^{-1/2})$ (dado que \hat{q}_{21} resulta de substituir em q_{21} as quantidades desconhecidas por estimativas *bootstrap*) então, a amplitude deste intervalo sofre de um erro $O_p(n^{-2})$. É de notar que este erro é inferior ao registado nos intervalos construídos com os outros métodos.

3.5 - DESENVOLVIMENTOS RECENTES

A teoria *bootstrap* de construção de I.C. mostra-se ainda em fase de expansão, surgindo frequentemente novos desenvolvimentos e perspectivas. É interessante referir três recentes abordagens que recorrem a processos iterativos com o objectivo de melhorar a aproximação entre a cobertura real e o nível nominal dos intervalos de confiança. São elas: o *bootstrap-t* modificado, introduzido por Peter Hall e Abramovitch e Singh; o método de calibragem de Loh; o método de pré-pivotação de Beran.

O *bootstrap-t* modificado recorre às expansões de Edgeworth da estatística *studentizada*. Para uma análise mais pormenorizada veja-se Abramovitch e Singh(1985), Peter Hall(1983) e Peter Hall(1986).

Suponha-se que a distribuição de $\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}$ verifica a E.E. (3.6). Se existir a respectiva expansão inversa de Cornish-Fisher (3.7) então é verdadeira a igualdade,

$$P\left\{\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma} \leq z_{\alpha} + \sum_{j=1}^v n^{-1/2} q_{j1}(z_{\alpha})\right\} = \alpha + O(n^{-(v+1)/2}) \quad (3.61)$$

O *bootstrap-t* modificado baseia-se nesta expressão para construir os I.C. com base na estatística *studentizada* modificada,

$$\sqrt{n} \frac{(\hat{\theta} - \theta)}{\hat{\sigma}} - \sum_{j=1}^v n^{-1/2} q_{j1}(z_{\alpha}) \quad (3.62)$$

A partir de (3.61) facilmente se nota que a estatística (3.62) permite uma aproximação à Normal com um erro de ordem inferior ao verificado pela aproximação (com um número igual de termos) da estatística não modificada. Portanto, regista-se nos I.C. com base em (3.62) um erro na cobertura inferior ao dos I.C. dados pelo *bootstrap-t*.

Geralmente como os q_{j1} não são conhecidos é necessário substituí-los pelas respectivas estimativas *bootstrap*. Como consequência, a expansão (3.62) falha para $v \geq 2$ (veja-se Hall(1986) p.1435). No entanto, a correcção com um termo conduz a I.C. unilaterais e bilaterais com um erro de cobertura de ordem $O_p(n^{-3/2})$, quando para os I.C. BCa e *bootstrap-t* era $O_p(n^{-1})$.

Este método é iterativo no sentido de que primeiro há que determinar a distribuição *bootstrap* da estatística *studentizada* de forma a obter a estimativa de q_{11} e, numa segunda fase definir a estatística modificada a partir da qual se processa a inferência. Hall(1983) apresenta uma variante deste processo onde se destaca ainda mais o carácter iterativo.

O método apresentado por Loh produz intervalos de confiança unilaterais que são equivalentes de segunda ordem ao *bootstrap-t* e ao BCa mas possuindo erros de cobertura de ordem inferior.

O método de calibragem aplica-se a qualquer método de construção de I.C. aproximados inclusivé ao *standard*. Caracteriza-se nos seguintes passos:

1) Em face de um dado I.C. aproximado ao nível α , estimar a sua cobertura real através das técnicas *bootstrap*. Esta estimativa é dada pela probabilidade dos intervalos gerados através de reamostragens a partir de \hat{F} conterem $\hat{\theta} = \theta(\hat{F})$. Loh(1987) prova para diversos métodos de construção de I.C. aproximados (inclusivé o método dos percentis e BC) que esta estimativa converge para a cobertura real .

2) Conhecendo uma estimativa da cobertura real, $\hat{\alpha}$, utilizar esta informação para construir intervalos com cobertura nominal mais próxima da real. Para tal, aplicar o método de construção de I.C. escolhido inicialmente mas considerando um novo nível nominal de confiança, α_1 , de forma que o novo I.C. calibrado tenha um nível de confiança real mais próximo de α . O processo pode ser repetido sucessivamente até ser encontrada uma cobertura real desejada. No entanto, como refere o autor:¹

In general, it would be impractical to iterate the calibration process until $\hat{\gamma}_n$ (cobertura real estimada) converges to the desired level. I have found that often a one-step calibration plus linear interpolation is enough.

1 Loh(1987) p.158

Esta metodologia consiste em considerar o novo α_1 a partir da expressão:

$$\alpha_1 = \frac{\alpha^2}{\hat{\alpha}} \quad \hat{\alpha} \geq \alpha$$

$$\alpha_1 = \alpha + \frac{(1-\alpha)}{(1-\hat{\alpha})}(\alpha - \hat{\alpha}) \quad \hat{\alpha} < \alpha$$

Grande parte dos I.C. são construídos utilizando uma variável pivô. Os intervalos de Efron fazem-no processando uma transformação prévia no parâmetro original enquanto que o *bootstrap-t* fá-lo através da studentização. Mas, se esta variável não constitui na realidade uma variável pivô, a cobertura real dos I.C. pode distanciar-se significativamente do nível nominal desejado.

Beran pretende atenuar este problema implementando uma transformação sobre a variável fulcral de forma a determinar uma nova variável com propriedades mais próximas de uma variável pivô, denominando este método de pré-pivotação.

Suponha-se que se constrói um I.C. para θ através da variável $t_n(\theta)$ (que depende da amostra e de θ). A transformação operada pela pré-pivotação sobre $t_n(\theta)$, originando $t_{n,1}(\theta)$, define-se por

$$t_{n,1}(\theta) = \hat{H}_n\{t_n(\theta)\} ,$$

em que \hat{H}_n é a distribuição *bootstrap* de $t_n(\theta)$. A nova variável $t_{n,1}(\theta)$ tem como distribuição limite a uniforme, sendo muito natural que a sua distribuição dependa menos de F e como tal de θ . Assim, $t_{n,1}(\theta)$ terá propriedades que se aproximam mais de um pivô, permitindo diminuir a grandeza do erro da cobertura real. Iterações adicionais podem ajudar esta aproximação. Beran(1987a) mostra

para algumas situações que a pré-pivotação aproxima a cobertura real da cobertura nominal de forma superior aos métodos *bootstrap* usuais.

Este método tem como vantagens por um lado não necessitar de recorrer a expansões analíticas, e por outro lado, pode ser implementado utilizando as técnicas de Monte Carlo. Contudo, exige níveis de computação extremamente elevados dado que para além do *bootstrap* usual necessita de um novo *bootstrap* sobre cada uma das réplicas originadas com a reamostragem inicial. Beran¹, descreve com pormenor esta implementação.

3.6 - ALGUMAS CONSIDERAÇÕES

Dos métodos *bootstrap* apresentados destacam-se o método com correcção do enviesamento e aceleração, BCa, e o *bootstrap-t* por verificarem propriedades assintóticas superiores ao método *standard*, permitindo melhorar a qualidade da cobertura dos intervalos de confiança unilaterais, e a aproximação aos extremos do I.C. exacto. O *bootstrap-t* permite ainda diminuir o erro da amplitude dos I.C. bilaterais para $O_p(n^{-5/2})$.

Estes dois métodos traduzem duas opções distintas. O primeiro conduz a inferência a partir da distribuição *bootstrap* da estatística conjugando determinadas correcções analíticas, enquanto que o segundo pressupõe a studentização de $\hat{\theta}$, obtendo os I.C. directamente a partir dos percentis da distribuição *bootstrap* da estatística *studentizada* (passando depois para o domínio de θ).

Podem ser distinguidas dois tipos de motivações para o BCa. A primeira, supõe que existe uma transformação normalizadora, que se verifica aproximadamente, com desvio padrão que varia em

1 Beran(1987a) pp.460 e 461 .

função do parâmetro. Diccio e Tibshirani mostram que esta transformação existe com erro de grandeza $O(n^{-1})$. A correcção da aceleração, a , representa a variação relativa do desvio padrão da transformação em relação à variação do transformado do parâmetro por unidade de desvio padrão. Efron apresenta estimativas para uma aproximação desta constante, aproximação que não afecta as propriedades de segunda ordem dos I.C. BCa. Mostra que em problemas de parâmetro unidimensional¹ o BCa é correcto de segunda ordem.

A segunda motivação enquadra-se na abordagem de Peter Hall que recorre a E.E. e de C.F.. Esta abordagem exige que a distribuição *studentizada* da estatística verifique uma E.E. e os respectivos quantis uma expansão de C.F.. A constante de aceleração é fixada de forma que se consiga diminuir a grandeza do erro na cobertura dos I.C. unilaterais para $O(n^{-1})$, e os extremos sejam correctos de segunda ordem. Em ambas as perspectivas a correcção do enviesamento tem igual interpretação. No entanto, para os problemas paramétricos com a família exponencial multiparâmetros e os problemas não paramétricos com a estatística função suave de vectores de médias, as duas abordagens são coincidentes.

O aspecto menos atraente do BCa reside na determinação da estimativa da constante de aceleração. A abordagem de Hall pressupõe a aplicação da fórmula (3.54a). Esta estimativa necessita da E.E. da distribuição $\hat{R}(x)$. Dado que este desenvolvimento geralmente é complexo fica comprometida a determinação de a por esta via. Por outro lado, a fórmula apontada por Efron para os problemas paramétricos pressupõe cálculos algo sofisticados nem sempre fáceis de determinar. Para os problemas não paramétricos, a estimação de a (conforme Efron) envolve estimativas das funções empíricas de influência. O ideal será considerar na sua estimação valores infinitesimais para ϵ mas usualmente opta-se pela forma muito mais operacional dada por (3.47), só que, ainda não estão devidamente estudadas as

1 Com $\hat{\theta}$ tendo um comportamento assintótico igual ao estimador M.V..

consequências deste procedimento na precisão (mais geralmente nas propriedades de segunda ordem) do BCa. As expressões apontadas por Babu e Konishi para a correcção do enviesamento e para a parecem de difícil implementação, a sua virtude reside muito mais no campo teórico permitindo assinalar as propriedades assintóticas do BCa.

Para os problemas (paramétricos e não paramétricos) em que a estatística e o seu desvio padrão são funções suaves de médias (multidimensionais) da amostra (problemas que, como já se referiu, englobam um vasto conjunto de casos) a estimação de σ não constitui problema pois aplica-se a fórmula operacional e directa de Hall (3.54).

Apesar de não necessitar de correcções analíticas, o *bootstrap-t* apresenta também alguns inconvenientes. A sua aplicação está dependente da existência de uma estimativa estável do desvio padrão da estatística. Em muitos problemas não paramétricos tal não existe e estimativas utilitárias como a jackknife são desaconselhadas. Por outro lado, o BCa detém propriedades de invariância que o *bootstrap-t* não possui.

O maior inconveniente do *bootstrap-t* reside na falha de funcionamento para alguns problemas, necessitando de uma transformação prévia. É o caso do coeficiente de correlação¹.

Motivado por esta vulnerabilidade do *bootstrap-t* Tibshirani(1988) sugere uma modificação. Como o *bootstrap-t* não é invariante sob reparametrização há que estabelecer uma transformação que estabilize a variância. Surge assim, o *bootstrap-t* com estabilização da variância. Num primeiro passo

1 Veja-se Tibshirani(1988) p. 443.

estima-se através do *bootstrap* uma transformação estabilizadora da variância assintótica, seja $h^*(\theta)$. Esta será a estimativa de,

$$\beta = h(\theta) = \int^{\theta} \sigma_s^{-1} ds ,$$

com σ_s o desvio padrão assintótico de $\hat{\theta}$ quando $\theta = s$. É de notar que esta transformação estabilizadora não é determinada a partir da família menos favorável de Stein como no $BC\alpha^0$ mas directamente com o *bootstrap*. Depois calcula-se um I.C. para $h^*(\theta)$. Por fim, passa-se desta I.C. para um intervalo no domínio de θ aplicando a transformação inversa. O intervalo obtido é invariante. O autor¹ descreve com pormenor o algoritmo que implementa este procedimento. Tem como inconveniente necessitar de níveis mais elevados de computação, com cálculos sofisticados.

Esta abordagem requer que a variância de $\hat{\theta}$ possa ser escrita como função de θ para além de constituir uma função conhecida de médias observáveis. Tibshirani(1988) refere que conserva as propriedades de segunda ordem do *bootstrap-t*.

Comparando os dois métodos seleccionados é de notar que o *bootstrap-t* apresenta algumas vantagens relativamente ao BCa. Teoricamente parece mais correcto por ser a estimativa *bootstrap* directa do extremo exacto, enquanto que o BCa produz a estimativa *bootstrap* de um extremo incorrecto, necessitando de correcções analíticas para se aproximar do verdadeiro extremo. Ainda relativamente a este assunto veja-se o exemplo de Peter Hall(1988) p.941, que trata um problema trivial, para o qual não é necessário o *bootstrap*, mas é interessante para comparar o comportamento das duas metodologias (BCa e *bootstrap-t*). Neste problema, a

¹ Tibshirani(1988) p. 437.

estimativa BCa é incorrecta (coincide com $\hat{\theta}_{\text{back}}(\alpha)$), enquanto que $\hat{\theta}(\alpha)$ consegue identificar-se com o verdadeiro extremo, o que mostra que as correcções analíticas do BCa nem sempre funcionam. A necessidade do uso de correcções analíticas não está totalmente de acordo com a filosofia do *bootstrap* pois por princípio recorre-se à computação intensiva para evitar este tipo de correcções, favorecendo este argumento a óptica do *bootstrap-t*. Por outro lado, (sendo $\hat{\sigma}$ escolhido correctamente), o *bootstrap-t* capta melhor as propriedades de terceira ordem dos extremos exactos.

Do exposto conclui-se que a teoria *bootstrap* de construção de I.C. dotou os estatísticos com poderosas alternativas ao método clássico de construção de intervalos aproximados. Estes métodos revelam-se computacionalmente mais exigentes mas em contrapartida apresentam propriedades assintóticas superiores ao método *standard*.

4 - O *BOOTSTRAP* NA REGRESSÃO

Em problemas de regressão o *bootstrap* reveste-se de formas específicas que necessitam de ser esclarecidas por não estarem contempladas na metodologia geral exposta no capítulo 2.

Neste capítulo, a aplicação do *bootstrap* ao modelo clássico de regressão linear múltipla, pressupondo que as variáveis residuais têm distribuição desconhecida, assume um certo destaque. Este problema permite de uma forma simples e clara ilustrar o funcionamento da lógica *bootstrap* na regressão. Por outro lado, de um ponto de vista formal, será interessante averiguar se o *bootstrap* consegue alcançar a performance dos métodos clássicos. Apesar de, neste problema, o *bootstrap* ser desnecessário na estimação, visto os métodos existentes serem eficientes, é fundamental na inferência estatística do modelo (pois com os métodos clássicos só é possível aplicar resultados assintóticos). Por fim, será também considerada a expressão particular do *bootstrap* em modelos dinâmicos autoregressivos.

4.1 - CARACTERIZAÇÃO DO MODELO

O modelo clássico de regressão linear múltipla, verifica as seguintes hipóteses:

$$\text{HIPÓTESE 1} - Y = X\beta + U,$$

onde $Y = [Y_i]$ é um vector ($n \times 1$) em que Y_i ($i=1, \dots, n$) são variáveis aleatórias observáveis sem erro; $X = [x_{ij}]$ é uma matriz ($n \times k$) de variáveis não aleatórias observáveis sem erro (variáveis explicativas); $\beta = [\beta_j]$ é um vector de k parâmetros desconhecidos,

fixos e não sujeitos a restrições; e finalmente $U = [U_i]$ é um vector (nx1) de variáveis aleatórias, variáveis residuais, não observáveis. Considerando apenas a observação i , a equação apresentada vem igual a,

$$Y_i = X_i \beta + U_i , \quad (4.1)$$

onde X_i representa a linha i da matriz X , i.e., a observação i de todas as variáveis explicativas.

$$\text{HIPÓTESE 2} - E\{U\} = \underline{0} .$$

$$\text{HIPÓTESE 3} - V\{U\} = \sigma^2 I_n ,$$

donde as variáveis residuais $U_i (i=1, \dots, n)$ são homocedásticas (com variância desconhecida σ^2) e não correlacionadas.

$$\text{HIPÓTESE 4} - U_i \sim F \text{ desconhecida}, (i=1, \dots, n)$$

$$\text{HIPÓTESE 5} - r(X) = k ,$$

com $r(X)$ a característica da matriz X .

O modelo dinâmico autoregressivo que irá ser objecto de análise deste capítulo, difere do modelo exposto apenas na constituição da matriz X . Assim, verifica todas as hipóteses apresentadas, considerando apenas na hipótese 1 que a matriz X inclui valores desfasados da variável endógena (para além dos valores das variáveis não aleatórias).

Os parâmetros desconhecidos, β e σ^2 , são estimados (em ambos os modelos) através do método dos Mínimos Quadrados (MQ). O estimador para β é igual a,

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (4.2)$$

Enquanto que o estimador para σ^2 é dado por,

$$\hat{\sigma}^2 = \frac{\hat{U}'\hat{U}}{n-k} \quad (4.3)$$

onde o vector $(n \times 1)$ \hat{U} contém os n resíduos do modelo, obtidos de acordo com,

$$\hat{U}_i = Y_i - X_i \hat{\beta} \quad (i=1, \dots, n) \quad (4.4)$$

4.2 - ALGORITMO *BOOTSTRAP*

A aplicação do *bootstrap* ao modelo de regressão definido pelas hipóteses 1 a 5, será abordada apenas na perspectiva numérica através da implementação do algoritmo de simulação de Monte Carlo, pois a dedução analítica das estimativas *bootstrap* em modelos econométricos é demasiado complexa, não sendo por isso prosseguida na generalidade dos problemas.

A ideia base do algoritmo *bootstrap* consiste em proceder a sucessivas reamostragens dos resíduos preservando a estrutura estocástica previamente estipulada para as variáveis residuais. Este algoritmo descreve-se nos seguintes passos:

- 1) Sejam os n resíduos do modelo obtidos de acordo com (4.4), $\hat{U}_1, \dots, \hat{U}_i, \dots, \hat{U}_n$. No algoritmo é pressuposto que os resíduos têm média nula (em conformidade com a

hipótese 2 sobre as variáveis residuais). No entanto, se o modelo não apresentar termo autónomo é muito possível que a média dos resíduos seja não nula, sendo necessário neste caso centrar previamente os resíduos subtraindo-lhes a média. A partir do conjunto dos n resíduos (centrados) construir a respectiva função de distribuição empírica, \hat{F} , que atribui probabilidade $1/n$ a cada \hat{U}_i , ($i=1, \dots, n$)!

Para cada reamostragem b :

- 2) Gerar aleatoriamente a partir de \hat{F} n valores que constituem o vector de resíduos *bootstrap*, \hat{U}^{*b} .
- 3) Com os resíduos *bootstrap* construir o vector das n observações artificiais, para a variável endógena, y^{*b} , através de,

$$y^{*b} = X\hat{\beta} + \hat{U}^{*b} \quad (4.5)$$

com $\hat{\beta}$ dado por (4.2).

- 4) Utilizando os dados y^{*b} e X , estimar os parâmetros do modelo pelo processo escolhido inicialmente. As k estimativas para β constituirão o vector $\hat{\beta}^{*b}$ e a estimativa de σ^2 será $\hat{\sigma}^{*b}$.
- 5) Repetir os pontos 2), 3) e 4) um número suficientemente grande, B , de vezes.

¹ Rayner(1988) refere uma alternativa diferente de reamostragem onde considera como fonte de reamostragem o conjunto dos resíduos mais os seus simétricos. Necessariamente este conjunto de valores tem média nula, não havendo necessidade de uma centragem prévia. O autor refere ainda que este *bootstrap* é válido em grandes amostras mesmo quando a distribuição das variáveis residuais não é simétrica, sendo assim assintoticamente robusto.

Desta forma criam-se B valores simulados para o estimador $\hat{\beta} = (X'X)^{-1}X'Y^*$. As estimativas *bootstrap* do enviesamento de $\hat{\beta}$, do desvio padrão de $\hat{\beta}$, e da distribuição de $\hat{\beta}$, são obtidas da forma apontada no algoritmo geral *bootstrap* (ponto 2.1) para uma estatística genérica $\hat{\theta}$. Como tal, para obter estas estimativas há apenas que substituir respectivamente em (2.2), (2.3), (2.4) e (2.5) $\hat{\theta}^{*b}$ por $\hat{\beta}^{*b}$.

As reamostragens alternativas mencionadas no ponto 2.2 podem conduzir a melhores resultados *bootstrap* na regressão, pelos mesmos motivos indicados na altura. Para as implementar há apenas que adaptar o passo 2 do algoritmo descrito. Assim, no caso do *bootstrap* alisado e do *bootstrap* paramétrico o vector dos n resíduos artificiais, U^{*b} , será gerado a partir de n realizações independentes da distribuição F , escolhida de acordo com os moldes desenvolvidos em 2.2.

Variante para modelos dinâmicos: Para o modelo que inclui valores desfasados da variável endógena, os pontos 3) e 4) do algoritmo *bootstrap* necessitam de um esclarecimento adicional. Suponha-se (sem prejuízo de generalização) que as observações das variáveis explicativas no período i, representadas por X_{it} , verificam $X_{it} = [Y_{i-1} \dots Y_{i-j} \ x_{ij+1} \dots x_{ik}]$. Admite-se que os valores Y_0, \dots, Y_{1-j} são conhecidos.

Para a reamostragem b, os dados *bootstrap* correspondentes à i-ésima observação são calculados de acordo com,

$$Y_i^* = [Y_{i-1}^* \dots Y_{i-j}^* \ x_{ij+1} \dots x_{ik}] \hat{\beta} + \tilde{U}_i^*, \quad (4.6)$$

em que \tilde{U}_i^* é o elemento i do vector de resíduos *bootstrap* determinado em 2). Para obter as n observações repete-se (4.6) iterativamente de $i=1$ a $i=n$, fazendo nas j primeiras observações $Y_{i-l}^* = Y_{i-l}$ ($l=0, 1, \dots, j-1$).

Os parâmetros do modelo são estimados através de (4.2) com os dados Y^* e X^* , em que X^* inclui os valores desfasados de Y^* para além das observações fixas das variáveis exógenas.

4.3 - PROPRIEDADES ASSIMPTÓTICAS

Considere-se ainda o modelo apresentado no ponto 4.1,

$$Y_n = X_n \beta + U_n ,$$

verificando as hipóteses 1 a 5. Suponha-se que β é estimado de acordo com (4.2). Os resultados clássicos estipulam que β_n tem média β e matriz de covariâncias $\sigma^2 \{X_n' X_n\}^{-1}$, enquanto que a distribuição assintótica de $\sqrt{n}(\beta_n - \beta)$ é uma Normal k -dimensional de média zero e matriz de covariâncias $\sigma^2 V^{-1}$, sendo $V = \text{plim}(1/n) X_n' X_n$, positiva definida. Pretende-se mostrar que o *bootstrap* conduz a resultados assintóticos coincidentes com os clássicos quando o processo de simulação é efectuado de forma consistente, respeitando as hipóteses fixadas para as variáveis residuais.

O estimador *bootstrap* dos MQ é igual a,

$$\hat{\beta}_n^* = (X_n' X_n)^{-1} X_n' Y_n^* .$$

Analise-se qual o valor esperado e qual a matriz de covariâncias em pequenas amostras deste estimador, quando as variáveis residuais *bootstrap*, U^* , tal como as variáveis U verificam as hipóteses 2 e 3,

$$E.\{\hat{\beta}^*\} = \hat{\beta} + (X'X)^{-1} X'E.\{U^*\} = \hat{\beta} ,$$

(4.7)

$$V.\{\hat{\beta}^*\} = (X'X)^{-1} X'E.\{U^{*'}U^*\}X(X'X)^{-1} = \sigma^2(X'X)^{-1} .$$

Nesta situação, a aplicação do *bootstrap* conduz aos resultados da estimação clássica.

Freedman(1981) demonstra as seguintes propriedades assintóticas do estimador dos *MQ bootstrap* no modelo definido,

- 1) A distribuição condicional de $\sqrt{n}(\hat{\beta}_n^* - \hat{\beta}_n)$ converge fracamente para a normal $N(0, \sigma^2 V^{-1})$, (com V conforme definido anteriormente),
- 2) A distribuição condicional do desvio padrão de $U_{1,n}^*$, $\hat{\sigma}_n^*$, (com $\hat{\sigma}_n^{*2} = (1/n) \sum U_{1,n}^{*2}$), converge para a distribuição que atribui probabilidade 1 a σ .
- 3) A distribuição condicional do pivô $\{X_{1,n}'X_n\}^{1/2} \{\hat{\beta}_n^* - \hat{\beta}_n\} / \hat{\sigma}_n^*$ converge fracamente para a Normal estandardizada k -dimensional.

As propriedades assintóticas enunciadas verificadas pelo *bootstrap* coincidem com as propriedades assintóticas referidas do estimador *MQ* clássico apresentado, denotando a validade do *bootstrap* em grandes amostras para este problema.

Freedman(1981) afirma ainda que a aproximação numérica *bootstrap* à distribuição de $\sqrt{n}(\hat{\beta}_n^* - \hat{\beta}_n)$ constitui uma boa aproximação da distribuição de $\sqrt{n}(\hat{\beta}_n - \beta)$ desde que n seja grande e $\sigma^2 k \text{tr}(X'X)^{-1}$ seja pequeno. No entanto, se os resíduos não forem centrados (quando não existe constante na regressão) o *bootstrap* geralmente falha. Isto porque, neste caso, a distribuição de $\sqrt{n}(\hat{\beta}_n^* - \hat{\beta}_n)$

incorpora um termo aleatório que traduz o enviesamento e depende das variáveis residuais, tendo geralmente uma distribuição limite Normal não degenerada.

É de esperar que para o modelo dinâmico autoregressivo estas propriedades assintóticas sejam ainda válidas, dado que verifica $\text{plim}(1/n)X_n'U_n = 0$, $\text{plim}\beta_n = \beta$, e $\text{plim}X_n'X_n = V$, donde se generaliza facilmente, para este caso, a demonstração feita para o modelo clássico.

Analise-se o comportamento assintótico das estimativas *bootstrap* da variância e do enviesamento de uma função do estimador MQ em grandes amostras, seja $\hat{\theta} = g(\beta)$, com g uma função de \mathbb{R}^k em \mathbb{R} .

Shao(1988b) aborda este assunto para um modelo geral que inclui o modelo clássico e o modelo dinâmico caracterizado em 4.1 como casos particulares, e demonstra que o enviesamento *bootstrap* e a variância *bootstrap* de $\hat{\theta}$ verificam, sob determinadas condições,¹

$$\xi_*(\hat{\theta}) = \xi(\hat{\theta}) + o_p(n^{-1}) , \quad (4.8)$$

$$V_*(\hat{\theta}) = V\{\hat{\theta}\} + o_p(n^{-1}) ,$$

onde $\beta(\hat{\theta})$ e $V\{\hat{\theta}\}$ representam respectivamente o verdadeiro enviesamento e a verdadeira variância de $\hat{\theta}$. Deduz também os seguintes limites superiores para $\xi_*(\hat{\theta})$ e $V_*(\hat{\theta})$,

¹ Estas condições consignam algumas condições de regularidade sobre g que são verificadas desde que a terceira e quarta derivadas de g sejam limitadas no espaço dos parâmetros. Shao propõe uma modificação do *bootstrap* a utilizar em situações não lineares quando as derivadas de g são complicadas ou não verificam as condições de regularidade apontadas. Para mais pormenores sobre este procedimento veja-se Shao(1988b) pp. 376 e 377.

$$E\{\xi.(\hat{\theta})\} = \xi(\hat{\theta}) + O(n^{-3/2}h_n^{1/2}) ,$$

(4.9)

$$E\{V.(\hat{\theta})\} = V(\hat{\theta}) + O(n^{-3/2}h_n^{1/2}) ,$$

com $h_n = \max_{1 \leq n} w_i$, em que $w_i = X_i'(X'X)^{-1}X_i$. As expressões em (4.9) mostram que o *bootstrap* estima de forma centrada o enviesamento e a variância de $\hat{\theta}$ até à ordem $O(n^{-3/2}h_n^{1/2})$. Do mesmo modo, o estimador *bootstrap* $\hat{\theta} - \xi.(\hat{\theta})$ é centrado para θ até à ordem mencionada.

Por sua vez, o erro quadrático médio de $\xi.(\hat{\theta})$ e de $V.(\hat{\theta})$ verifica,

$$MSE[\xi.(\hat{\theta})] = O(n^{-2}h_n) , \quad MSE[V.(\hat{\theta})] = O(n^{-2}h_n) ,$$

significando que quando $h_n \rightarrow 0$ $\xi.(\hat{\theta})$ e $V.(\hat{\theta})$ são consistentes no sentido que $\xi.(\hat{\theta}) = o(n^{-2})$ e $V.(\hat{\theta}) = o(n^{-2})$.

Em modelos autoregressivos Bose(1988) analisa a convergência da distribuição *bootstrap studentizada* do estimador MQ. Supõem-se que o modelo em questão é estacionário, sendo igual a,

$$Y_t = \sum_{i=1}^p \beta_i Y_{t-i} + U_t ,$$

em que as variáveis residuais U_t ($t=1, \dots, n$) verificam as hipóteses 2-4 expressas em 4.1.

O autor demonstra que, verificando-se determinadas condições, para quase todos os (Y_i) é válida a expressão,

$$\sup_x |P.(\sqrt{n} \Sigma_n^{-1/2} (\beta_n^* - \beta) \leq x) - P(\sqrt{n} \Sigma^{-1/2} (\beta_n - \beta) \leq x)| = o(n^{-1/2}), \quad (4.10)$$

onde Σ_n^* e Σ representam, respectivamente, a matriz de covariâncias de β^* e β . A partir de (4.10) fica patente que a distribuição *bootstrap studentizada* constitui uma aproximação com erro inferior à aproximação dada pela Normal standardizada.

4.4 - ALGUMAS CONSIDERAÇÕES

A estimação em modelos dinâmicos constitui um dos problemas para os quais a teoria econométrica ainda não está completamente estabelecida. Essencialmente porque não são conhecidas as propriedades estatísticas em pequenas amostras dos estimadores. Como tal, este problema pode constituir um exemplo em que aplicação do *bootstrap* pode ser mais proveitosa do que a utilização dos resultados assintóticos clássicos existentes. A vantagem do *bootstrap* pode ser manifesta nomeadamente na estimação da distribuição de β (veja-se o resultado de Bose sobre a superioridade da distribuição *bootstrap studentizada*) e na determinação de I.C. através dos métodos expostos no capítulo 3.

A verificação das propriedades do *bootstrap* deduzidas tem implícita a vigência das hipóteses identificadas para as variáveis residuais, nomeadamente serem identicamente distribuídas. Esta hipótese é violada quando existe heteroescedasticidade, i.e., as variáveis aleatórias U_i ($i=1, \dots, n$) têm desvio padrão diferente. Nesta situação, Wu(1986) e Shao(1988a) apontam que a estimativa

bootstrap (dos MQ) é geralmente enviesada e inconsistente. Este facto explica-se porque o processo de reamostragem descrito no passo 1) do algoritmo *bootstrap* pressupõem a situação *i.i.d.* perdendo-se assim a heterogeneidade dos resíduos. Por seu turno, Shao(1988a) mostra que o estimador *bootstrap* do enviesamento de β é inconsistente em presença de heteroscedasticidade, quando existem observações influentes no modelo. É de notar que os autores consideram o estimador (4.2) que não é eficiente na presença de heteroscedasticidade. Nesta situação Wu(1986) defende a aplicação de variantes do *jackknife*, nomeadamente o *jackknife* ponderado onde se eliminam de cada vez r observações, e uma variante do *bootstrap* que será referida no próximo parágrafo. Estes métodos, que o autor demonstra serem consistentes, têm a vantagem de se mostrarem robustos com o estimador Ordinário dos Mínimos Quadrados, evitando formas de estimação muito mais complexas como o método Generalizado dos Mínimos Quadrados.

Na estrutura de reamostragem *bootstrap* avançada por Wu(1986) são geradas observações *i.i.d.* a partir de uma população externa, não relacionada com o conjunto original dos dados. Para obter as observações artificiais primeiro há que ponderar os resíduos por um factor assintoticamente igual à unidade. Assim, a partir do resíduo \hat{U}_i ($i=1, \dots, n$) obtém-se o resíduo ponderado $\hat{V}_i = \hat{U}_i / \sqrt{1-w_i}$ com $w_i = X_i'(X'X)^{-1}X_i$. De seguida obtêm-se as observações *bootstrap* de acordo com,

$$Y_i^* = X_i \hat{\beta}' + \hat{V}_i t_i \quad (i=1, \dots, n) ,$$

em que t_i é uma variável aleatória de média nula e variância igual à unidade. Regina Liu¹ mostra que considerando o terceiro momento central de t_i igual a 1, para além dos requisitos mencionados, então o *bootstrap* de Wu verifica a propriedade

¹ Regina Liu(1988). Bootstrap procedures under some non-I.I.D. models. *The Annals of Statistics*, 16, 1696-1708.

assimptótica de segunda ordem do *bootstrap* clássico consistindo na correcção total do segundo termo da E.E. da distribuição *studentizada* de β^* .

Efron e Gong(1983) definem uma reamostragem alternativa, determinada a partir das observações (Y_i, X_i) .¹ Com este procedimento, já não existe uma igualdade em pequenas amostras entre os resultados *bootstrap* e os resultados tradicionais no modelo clássico. Os estimadores *bootstrap* são geralmente enviesados e inconsistentes, quer com erros homoescedásticos quer heteroescedásticos. No entanto, Freedman(1981) analisa as propriedades assimptóticas deste *bootstrap* para um modelo onde as variáveis explicativas são aleatórias, sendo os n vectores de observações $[Y_i, X_i]$ i.i.d. e $E\{X_{ij}, U_i\} = 0 \quad j=1, \dots, k$. Esta formulação inclui a hipótese de heteroescedasticidade. O autor mostra que a distribuição assimptótica de $\sqrt{n}(\hat{\beta}_n^* - \beta_n)$ tende em lei para a normal com média 0 e matriz de covariâncias $V^{-1} M V^{-1}$, com M uma matriz em que $M_{jk} = E\{X_{ij} X_{ik} U_i^2\}$. Este resultado mostra que o *bootstrap* a partir das observações é assimptoticamente correcto para o modelo referido. Por outro lado, este método depende menos da estrutura da regressão, podendo como tal fornecer uma estimativa fiável da variabilidade de β mesmo quando o modelo não está explicitado correctamente.

A Econometria apresenta ainda um campo vasto de problemas onde a aplicação do *bootstrap* pode ajudar a encontrar soluções meritórias. Para além da (já referida) estimação em modelos dinâmicos salientam-se ainda os problemas de heteroescedasticidade (sobre os quais se fez uma breve abordagem),

1 Veja-se Efron e Gong(1983) p.43

autocorrelação¹, alguns problemas de previsão, entre outros. É de realçar ainda a área de ensaios de hipóteses, nomeadamente certos testes sobre a quebra da estrutura da regressão², e de uma forma geral, todos os testes onde não são conhecidas as distribuições em pequenas amostras. Todos estes assuntos merecem uma investigação mais alargada.

Na aplicação do *bootstrap* convém não esquecer que a consistência do método não é robusta em relação às hipóteses que especificam o modelo. Como tal, no processo de reamostragem há que preservar a estrutura probabilística das variáveis residuais (com heteroscedasticidade é aconselhado o procedimento de Wu). Para os casos em que tal não é possível (até por haver uma grande incerteza sobre a estrutura do modelo) é preferível optar pela reamostragem das observações.

1 Sobre este assunto menciona-se o texto de D. Prescott e T. Stengos (1987). Hypothesis testing in regression models with AR(1) errors and a lagged dependent variable. Bootstrapping and artificial regression compared. *Economics Letters*, 24, 237-242; e o texto, Veall, M.R. (1986). Bootstrapping regression estimators under first-order serial correlation. *Economics Letters*, 21, 41-44.

2 Christiano, L.J. (1988); Searching for a break in GNP, *NBER working paper*, #2695; desenvolve uma aplicação deste tipo.

5 - APLICAÇÃO EMPÍRICA: ESTIMAÇÃO E PREVISÃO DO CONSUMO

Este capítulo tem como objectivo a aplicação empírica das técnicas *bootstrap* apresentadas, a um problema de estimação e previsão com um modelo dinâmico autoregressivo.

Escolheu-se o Consumo como tema desta aplicação empírica dado que nos estudos económicos empíricos a modelização deste agregado constitui um dos exemplos mais frequentes de regressão com variável endógena desfasada.

Não foi propósito elaborar um estudo sobre a especificação empírica ideal da função consumo até porque sobre este tema já existe o excelente trabalho de Santos Silva(1988). O objectivo principal desta aplicação foi antes de tudo ilustrar a *performance* das técnicas *bootstrap*. Como tal, procurou-se antes uma modelização que obedecesse a um conjunto de critérios decorrentes das condições restrictivas desta aplicação. Por um lado houve que limitar o número de variáveis explicativas para que a aplicação empírica fosse compoprável. Isto porque, para além do *bootstrap* ser bastante exigente computacionalmente, a previsão em períodos com dados desconhecidos incorreu na necessidade de se proceder à modelização empírica das variáveis exógenas aplicando simultaneamente o *bootstrap* a estes ajustamentos. Por outro lado, pretendia-se uma modelização onde a evidência empírica não rejeitasse as hipóteses clássicas do modelo de regressão caracterizado em 4.1. Foi particularmente difícil compatibilizar a necessidade de parcimónia com a inexistência de autocorrelação nos resíduos, dado que esta costuma espelhar uma insuficiência da especificação do modelo. No entanto foi possível chegar à equação que explica a taxa de crescimento do Consumo em função da taxa de crescimento do Rendimento no período anterior e do Consumo desfasado, de acordo com,

$$DC_t = \beta_1 + \beta_2 DY_{t-1} + \beta_3 c_{t-2} + U_t ,$$

$$DC_t = c_t - c_{t-1} , \quad (5.1)$$

$$DY_{t-1} = y_{t-1} - y_{t-2} ,$$

onde c_t corresponde ao logaritmo natural das Despesas de Consumo a preços constantes de 1977 e y_t refere-se ao logaritmo natural do Rendimento Disponível a preços de 1977¹. Os dados foram extraídos de Cartaxo e Rosa(1986) e referem-se ao período de 1958 a 1985 (inclusivé).

No decorrer do capítulo serão determinadas estimativas do enviesamento de β_1 , β_2 , β_3 e $\hat{\sigma}$ (obtidos com o *MOMQ*), I.C. aproximados para β_1 , β_2 , β_3 . No domínio da previsão, serão calculadas as previsões para o Consumo em 1986 e 1987 assim como estimativas *bootstrap* da média e desvio padrão do erro de previsão, e I.C. aproximados para os valores previstos.

A computação foi feita com os meios informáticos disponíveis no Centro Informático do I.S.E.G., recorrendo ao *package* TSP (a estimação de funções de densidade espectral foi feita com o *package* GENSTAT). Na posse dos B valores simulados, os I.C. *bootstrap* foram determinados com base numa adaptação feita ao programa em Pascal de Andrade e Silva incluído em Murteira(1988).

5.1 - O *BOOTSTRAP* NA REGRESSÃO

5.1.1 - Estimativas clássicas

A estimação de (5.1) através do Método Ordinário dos Mínimos Quadrados (*MOMQ*) conduziu aos seguintes resultados,

¹ Este foi obtido aplicando ao Rendimento Disponível a preços constantes o deflator do Consumo patente em Cartaxo e Rosa(1986).

QUADRO 1 - ESTIMATIVAS DO MOMQ

coef.	estim.	desv.pad.	est. t
β_1	0.47796	0.18251	2.62
β_2	0.40551	0.13792	2.94
β_3	-0.03630	0.01423	-2.55
	$R^2 = 0.52$	$\bar{R}^2 = 0.48$	$F_{(2,23)} = 12.6$

As estimativas dos parâmetros podem considerar-se estatisticamente não nulas (pois os respectivos quocientes t são claramente superiores a 2). O sinal da estimativa do coeficiente de DY_{t-1} não constitui surpresa, pois espera-se que um incremento da taxa de crescimento do rendimento no período anterior tenha repercussão positiva no crescimento do consumo presente. Influência oposta tem a variação do nível de consumo em t-2. De acordo com β_3 , quanto maior a satisfação dos agentes em t-2 menor a necessidade de consumo presente, embora esta influência seja de muito fraca magnitude.

O valor da estatística F permite concluir que as variáveis explicativas consideradas são determinantes, no seu conjunto, para a explicação da variação da taxa de crescimento do consumo. No entanto, o coeficiente de determinação não é muito elevado, reflectindo a ausência de variáveis que tradicionalmente explicam o andamento do consumo, como a taxa de juro e a taxa de inflação. Contudo o valor apresentado é já razoável.

5.1.2 - Análise estatística da especificação empírica

Com o intuito de obter uma avaliação mais completa da especificação empírica do modelo procedeu-se à aplicação de um conjunto de testes de diagnóstico. Com estes testes pretendeu-se detectar algumas perturbações que afectam a eficiência do *MOMQ*, como a existência de observações influentes, de autocorrelação e heteroescedasticidade. Foi também examinada a normalidade das variáveis residuais. Todos os testes utilizados estão descritos no anexo 1. Neste capítulo apenas se pretende mostrar e interpretar os resultados obtidos decorrentes da sua aplicação.

Observações influentes: Na detecção de observações influentes (*outliers*) foram analisados os índices w_i , $DFITS_i$, a função empírica de influência de cada β_j ($j=1,2,3$) e o teste $OUT(i)$. Os resultados obtidos constam do quadro 2 enquanto que as funções de influência empíricas estão representadas nas figuras 1 a 3.

1970	0.042	0.101	0.103
1971	0.050	-0.170	-0.109
1972	0.050	0.143	-0.015
1973	0.226	0.040	1.024
1974	0.000	0.041	0.000
1975	0.072	0.042	0.019
1976	0.077	0.069	0.280
1977	0.202	-0.073	-0.007
1978	0.010	-0.050	-0.036
1979	0.047	-0.041	-0.049
1980	0.057	0.047	0.162
1981	0.054	-0.120	-0.000
1982	0.129	0.016	0.001
1983	0.000	-0.000	-0.000

QUADRO 2 - OBS. INFLUENT.

i	w_t	$DFITS_t$	OUT(i)
1960	0.170	-0.330	-0.721
1961	0.160	0.366	0.834
1962	0.143	-1.284	-4.065
1963	0.121	0.339	0.910
1964	0.109	-0.084	-0.237
1965	0.088	0.195	0.621
1966	0.075	-0.074	-0.253
1967	0.072	0.045	0.159
1968	0.179	-0.188	-0.396
1969	0.067	0.233	0.866
1970	0.042	0.165	0.785
1971	0.058	-0.178	-0.709
1972	0.052	0.145	0.612
1973	0.225	0.990	1.948
1974	0.200	0.041	0.080
1975	0.072	0.247	0.879
1976	0.077	0.359	1.260
1977	0.069	-0.075	-0.268
1978	0.078	-0.555	-2.030
1979	0.097	-0.341	-1.043
1980	0.097	0.047	0.142
1981	0.094	-0.120	-0.363
1982	0.128	0.316	0.821
1983	0.090	-0.353	-1.131

1984	0.198	-0.556	-1.127
1985	0.242	0.587	1.042

FIG. 1 - F. INFLUÊNCIA DE $\hat{\beta}_1$

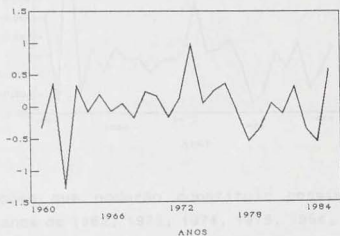


FIG. 2 - F. INFLUÊNCIA DE $\hat{\beta}_2$

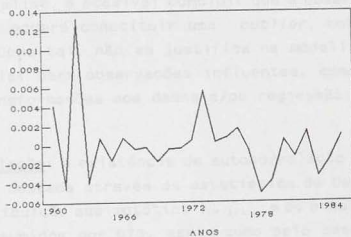
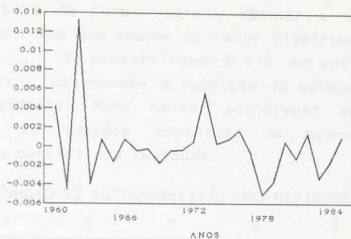


FIG. 3 - F. INFLUÊNCIA DE $\hat{\beta}_3$



As observações que poderão constituir possíveis *outliers* referem-se aos anos de 1962, 1973, 1974, 1978, 1984, 1985. Contudo os respectivos w_i são inferiores a $3K/N$. A exceção de OUT(62) todos os valores assumidos por OUT(i) são inferiores (em módulo) ao ponto crítico a 5% para esta estatística, 3.52. No entanto, o valor absoluto de OUT(62) é inferior ao ponto crítico a 1%, 4.18. Desta análise, é possível concluir que a observação de 1962 é a única que poderá constituir uma *outlier*, embora não seja muito claro. Como tal, não se justifica na modelização um tratamento especial para observações influentes, como a aplicação de certas transformações aos dados e/ou regressão robusta.

Autocorrelação: A existência de autocorrelação nas variáveis residuais foi testada através da estatística de Durbin de 1970, D70, com distribuição assintótica $t_{(N-k-2)}$, e do ensaio de Godfrey. Os valores assumidos por D70, assim como pelo teste de Godfrey contra as alternativas de autocorrelação de primeira ordem (GOD1), segunda ordem (GOD2), terceira ordem (GOD3) e quarta ordem (GOD4), foram,

D70	GOD1	GOD2	GOD3	GOD4
2.2	0.419	0.439	0.463	0.518

A estatística de Durbin parece indicar a existência de autocorrelação visto que assume um valor ligeiramente superior ao ponto da $t_{(N-k-2)}$ de probabilidade 0.975. No entanto, o teste de Godfrey conduz claramente à rejeição de autocorrelação das variáveis residuais. Para melhor esclarecer este assunto, analisaram-se as funções empíricas de autocorrelação e autocorrelação parcial dos resíduos.

FUNÇÃO DE AUTOCORRELAÇÃO DOS RESÍDUOS

t	1	2	3	4	5	6
$\rho(t)$	-0.127	0.027	-0.022	0.046	-0.083	0.104
	(-0.65)	(0.14)	(-0.11)	(0.23)	(-0.41)	(0.52)

FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL DOS RESÍDUOS

t	1	2	3	4	5	6
$\psi(t)$	-0.127	0.011	-0.017	0.042	-0.072	0.085
	(-0.65)	(0.06)	(-0.09)	(0.21)	(0.37)	(0.43)

Da análise do quociente t relativo às autocorrelações e autocorrelações parciais empíricas, é possível concluir que são estatisticamente nulas, conduzindo à rejeição da hipótese das variáveis residuais terem um comportamento AR(1) ou MA(1). A conclusão contrária obtida com o teste de Durbin pode ser explicada por a distribuição de D70 em pequenas amostras se desviar significativamente da sua distribuição assintótica.

Perante estes resultados, parece ser verosímil optar pela hipótese de inexistência de autocorrelação das variáveis residuais.

Heteroescedasticidade: Para averiguar uma possível existência de heteroescedasticidade nas variáveis residuais utilizou-se o teste de Engle que tem como alternativa a hipótese de se verificarem um processo com heteroescedasticidade condicional autoregressiva (ARCH), tendo sido testado sucessivamente de primeira (ENG1), segunda (ENG2), terceira (ENG3) e quarta (ENG4), ordens. Foram também utilizados o teste de White e o teste de Breush e Pagan, este último contra as alternativas de $\sigma^2 = h(\underline{Z}_t^1, \underline{\alpha})$ com $\underline{Z}_t^1 = (1, DY_{t-1}, c_{t-2})$, identificado por BP1, e $\sigma^2 = h(\underline{Z}_t^2, \underline{\alpha})$, $\underline{Z}_t^2 = (1, \log DY_{t-1}, \log c_{t-2})$, identificado por BP2. De seguida apresentam-se os valores obtidos para as estatísticas respectivas aos testes referidos.

ENG1	ENG2	ENG3	ENG4	WHITE	BP1	BP2
0.102	0.669	0.721	3.115	2.911	0.534	0.788

Os resultados indicados conduzem à aceitação da homoscedasticidade dos resíduos a um nível de significância de 10%.

Normalidade: A normalidade dos erros foi avaliada através da análise dos coeficientes empíricos de assimetria, $\sqrt{b_1}$, e kurtosis, b_2 , dos resíduos. Foi também analisada a estatística $BB = (N/26)[\sqrt{b_1}]^2 + (N/24)b_2^2$, assim como os testes de Shapiro-Wilk (SW), Shapiro-Francia (SF), e Kiefer e Salmon (KS). Obtiveram-se os seguintes resultados,

$\sqrt{b_1}$	b_2	BB	SW	SF	KS
-0.668	0.008	1.932	0.945	0.944	3.250

Todos estes valores conduzem à aceitação da hipótese, ao nível de significância de 10%, das variáveis residuais terem uma distribuição Normal.

Por fim, realizou-se o teste de Jarque e Bera que ensaia a hipótese conjunta dos erros serem simultaneamente não correla-

cionados, homoescedásticos, e com distribuição Normal. O valor obtido, 2.9, permite aceitar esta hipótese ao nível de significância de 10%.

Da aplicação do conjunto de testes de diagnóstico explicitado é possível concluir que a especificação empírica encontrada constitui uma modelização satisfatória do crescimento do consumo.

5.1.3 - Estimativas *Bootstrap*

Dado que não existem indícios de heteroescedasticidade e correlação temporal das variáveis residuais, nem sinais inequívocos de *outliers*, torna-se adequado aplicar o procedimento *bootstrap* descrito no capítulo 4, na variante para modelos dinâmicos.

Como foi aceite a hipótese das variáveis residuais terem distribuição Normal, também se implementou a versão paramétrica do algoritmo *bootstrap* (conforme o exposto em 2.2 e 4.1), para além do *bootstrap* convencional. No *bootstrap* paramétrico a reamostragem foi estruturada a partir de uma distribuição Normal com média nula e desvio padrão igual à estimativa do *MOMQ* para o desvio padrão dos erros.

Em primeiro lugar há que escolher o valor de B, número de iterações do algoritmo *bootstrap*. Como se pretende determinar intervalos de confiança o número mínimo aconselhado para B é 1000. Por outro lado, sabe-se que quanto maior B, maior é a qualidade das estimativas numéricas, exigindo em contrapartida custos de computação mais elevados. É necessário ponderar até onde o incremento destes custos é compensado pela variação que induzem nas estimativas.

Com base nesta ideia, seguiu-se uma metodologia heurística para a determinação de B. Procedeu-se ao *bootstrap* da equação (5.1) (com a reamostragem a partir da função de distribuição empírica dos resíduos) aumentando sucessivamente o valor de B, ponderando o tempo gasto em CPU e a variação das estimativas para os parâmetros e respectivos desvios padrões. Desta forma seleccionou-se o valor B=6000, pois para B=7000 a alteração das estimativas é insignificante não compensando o tempo gasto em CPU. Como tal, todos os resultados *bootstrap* apresentados daqui em diante foram obtidos com B=6000.

QUADRO 3 - ESTIM. BOOSTRAP DOS COEF. EM FUNÇÃO DE B

coef.	B=1000	B=5000	B=6000	B=7000
β_1	0.49389	0.500634	0.50102	0.50080
β_2	0.40763	0.40420	0.40339	0.40505
β_3	-0.03755	-0.03807	-0.03810	-0.03809

QUADRO 4 - ESTIM. BOOSTRAP DO D.P. DOS COEF. EM FUNÇÃO DE B

coef.	B=1000	B=5000	B=6000	B=7000
β_1	0.18096	0.18223	0.18338	0.18225
β_2	0.12592	0.12913	0.13007	0.12986
β_3	0.01415	0.01426	0.01423	0.01425

Sendo o modelo dinâmico as estimativas dadas pelo MOMQ não são centradas, conforme foi já referido. Com o *bootstrap* é possível determinar uma estimativa do seu enviesamento.

QUADRO 5 - ESTIM. *BOOT.* E *BOOT.* PARAM. DO ENV. DOS COEF.

COEF	MOMQ (1)	BOOT. (2)	ENVIES. (2)-(1)	B.PARAM. (4)	ENVIES. (4)-(1)
β_1	0.47796	0.50102	0.02306	0.49857	0.02061
β_2	0.40551	0.40339	-0.00212	0.40580	0.00029
β_3	-0.03630	-0.03810	-0.00180	-0.03792	-0.00162

QUADRO 6 - ESTIM. *BOOT.* E *BOOT.* PARAM. DO ENV. DO D.P. DOS COEF.

COEF	MOMQ (1)	BOOT. (2)	ENVIES. (2)-(1)	B.PARAM. (4)	ENVIES. (4)-(1)
β_1	0.18251	0.18202	-0.00049	0.18623	0.00372
β_2	0.13792	0.13007	-0.00785	0.13100	-0.00692
β_3	0.01423	0.01423	0.00000	0.01456	0.00033

As estimativas dadas pelo *bootstrap* paramétrico para os coeficientes estão mais próximas das obtidas com o MOMQ, enquanto que as estimativas do desvio padrão são ligeiramente superiores (facto que não é de admirar tendo em conta que a reamostragem a partir da distribuição Normal cria uma maior variabilidade para os erros simulados).

As estimativas *bootstrap* do enviesamento do estimador OMQ, quer para os β_j , $j=(1,2,3)$, quer para os respectivos desvios padrões, são muito reduzidas (é de notar que para o desvio padrão se considerou o enviesamento relativamente à estimativa *bootstrap*). Assim, neste problema, o enviesamento do estimador OMQ deve ser negligenciável, não havendo necessidade de centrar as respectivas estimativas.

5.1.4 - Intervalos de confiança *Bootstrap* para os coeficientes

Em modelos dinâmicos não é possível determinar I.C. exactos para os parâmetros pois a distribuição do estimador *OMQ* é desconhecida em pequenas amostras. Assim a distribuição $t_{(N-k)}$ para $(\beta_j - \hat{\beta}_j)/\hat{\sigma}_{\hat{\beta}_j}$ ($j=1,2,3$), é válida apenas assintoticamente.

O histograma dos 6000 valores *bootstrap* determinados constitui uma estimativa da função de densidade de probabilidade dos estimadores. Nas figuras 4, 5, e 6 estão representados os histogramas dos valores assumidos respectivamente por $\hat{\beta}_1$, $\hat{\beta}_2$, e $\hat{\beta}_3$, através do *bootstrap* convencional e do *bootstrap* paramétrico. É de notar uma grande regularidade na sua forma.

A partir das frequências relativas acumuladas dos valores *bootstrap* encontrados para os coeficientes construíram-se as distribuições empíricas para cada um dos $\hat{\beta}_j$ ($j=1,2,3$), assim como as respectivas distribuições empíricas *studentizadas*, patentes nas figuras 7 a 9. Existe uma semelhança quase total entre as distribuições resultantes do *bootstrap* convencional e do *bootstrap* paramétrico. É realçante a regularidade que denotam apresentando valores bastante reduzidos para o kurtosis e assimetria. Recorde-se o resultado de Bose em modelos autoregressivos, apontado no capítulo 4, sobre a superioridade da aproximação *bootstrap* à distribuição de $\hat{\beta}_j$ relativamente à Normal.

Será interessante analisar os percentis da distribuição empírica *studentizada* dos $\hat{\beta}_j$ ($j=1,2,3$) e compará-los com os percentis de igual ordem da $t_{(23)}$ teórica, valores incluídos no quadro 7.

QUADRO 7 - PERCENTIS DA DIST. *BOOTSTRAP STUDENTIZADA*

	a 2.5%		a 5%		a 95%		a 97.5%	
	B. con.	B. par.	B. con.	B. par.	B. con.	B. par.	B. con.	B. par.
β_1	-1.852	-1.903	-1.488	-1.540	1.759	1.751	2.040	2.087
β_2	-1.939	-1.933	-1.662	-1.626	1.621	1.643	1.923	1.978
β_3	-2.054	-2.066	-1.760	-1.731	1.478	1.539	1.864	1.903
t_{23}	-2.069		-1.714		1.714		2.069	

Nestes valores nota-se que as distribuições não são simétricas, embora a assimetria não deva ser muito importante. Relativamente à $t_{(23)}$, para β_1 há maior probabilidade em assumir valores mais elevados, enquanto que esta tendência é menos nítida em β_2 e particularmente em β_3 onde os percentis são muito próximos da $t_{(23)}$.

Na construção de I.C. aproximados para os parâmetros utilizaram-se as técnicas expostas no capítulo 3. Também foi determinado o I.C. clássico com base na distribuição assintótica $t_{(23)}$. Os resultados obtidos estão patentes nos quadros 8 a 10.

Convém precisar a forma utilizada para estimar a constante de aceleração. Optou-se por duas vias. Uma, dentro da perspectiva de Efron, consistindo em estimar a através de (3.46), com U_i a função de influência empírica de β_j ($j=1,2,3$) obtida em (3.47). A outra via decorre dos desenvolvimentos de Peter Hall, estimando o valor de a com (3.54) que para o j -ésimo coeficiente assume a forma,

$$\hat{a} = \frac{1}{6} \text{skew} \{ \sqrt{n}(\hat{\beta}_j^* - \hat{\beta}_j) \} ,$$

com $\text{skew} \{ \sqrt{n}(\hat{\beta}_j^* - \hat{\beta}_j) \}$ o coeficiente de assimetria da distribuição empírica dos 6000 valores $\sqrt{n}(\hat{\beta}_j^* - \hat{\beta}_j)$.

QUADRO 8 - INTERV. DE CONFIANÇA PARA β_1

	90%	95%
<i>com base na $t_{(23)}$</i>	0.16514 , 0.79078	0.10035 , 0.85557
<i>método percen.</i>		
<i>bootstrap</i>	0.20704 , 0.79818	0.14086 , 0.84932
<i>paramet. boots.</i>	0.19120 , 0.80413	0.12362 , 0.86663
<i>correc. envies.</i>		
<i>bootstrap</i>	0.15584 , 0.75412	0.09001 , 0.81240
<i>paramet. boots.</i>	0.15043 , 0.76810	0.08905 , 0.82049
<i>c.env/ac.-EFRON</i>		
<i>bootstrap</i>	0.10749 , 0.72848	0.01029 , 0.77538
<i>paramet. boots.</i>	0.10427 , 0.74213	0.03034 , 0.78776
<i>c.env./ac.-HALL</i>		
<i>bootstrap</i>	0.14843 , 0.74937	0.08494 , 0.80608
<i>paramet. boots.</i>	0.14412 , 0.76397	0.08654 , 0.81833
<i>bootstrap-t</i>		
<i>bootstrap</i>	0.15688 , 0.74962	0.10560 , 0.81598
<i>paramet. boots.</i>	0.15831 , 0.75900	0.09706 , 0.82523

QUADRO 9 - INTERV. DE CONFIANÇA PARA β_2

	90%	95%
<i>com base na $t_{(23)}$</i>	0.16912 , 0.64190	0.12015 , 0.69087
<i>método percen.</i>		
<i>bootstrap</i>	0.18933 , 0.61640	0.15334 , 0.65564
<i>paramet. boots.</i>	0.19252 , 0.62069	0.15233 , 0.66456
<i>correc. envies.</i>		
<i>bootstrap</i>	0.19442 , 0.62213	0.15798 , 0.66231
<i>paramet. boots.</i>	0.19252 , 0.62069	0.15229 , 0.66456
<i>c.env/ac.-EFRON</i>		
<i>bootstrap</i>	0.21230 , 0.64351	0.18103 , 0.69553
<i>paramet. boots.</i>	0.21069 , 0.63657	0.17511 , 0.69949
<i>c.env./ac.-HALL</i>		
<i>bootstrap</i>	0.19510 , 0.62239	0.15924 , 0.66267
<i>paramet. boots.</i>	0.19370 , 0.62197	0.15336 , 0.67101
<i>bootstrap-t</i>		
<i>bootstrap</i>	0.18189 , 0.63474	0.14027 , 0.67290
<i>paramet. boots.</i>	0.17896 , 0.62976	0.13277 , 0.67207

QUADRO 10 - INTERV. DE CONFIANÇA PARA β_3

	90%	95%
<i>com base na $t_{(23)}$</i>	-0.06069 , -0.01191	-0.06574 , -0.00686
<i>método percen.</i>		
<i>bootstrap</i>	-0.06134 , -0.01526	-0.06553 , -0.00978
<i>paramet. boots.</i>	-0.06150 , -0.01390	-0.06639 , -0.00860
<i>correc. envies.</i>		
<i>bootstrap</i>	-0.05808 , -0.01139	-0.06274 , -0.00639
<i>paramet. boots.</i>	-0.05862 , -0.01026	-0.06283 , -0.00599
<i>c.env/ac.-EFRON</i>		
<i>bootstrap</i>	-0.05646 , -0.00789	-0.05990 , -0.00055
<i>paramet. boots.</i>	-0.05680 , -0.00702	-0.06052 , -0.00107
<i>c.env./ac.-HALL</i>		
<i>bootstrap</i>	-0.05790 , -0.01113	-0.06237 , -0.00574
<i>paramet. boots.</i>	-0.05847 , -0.00990	-0.06258 , -0.00540
<i>bootstrap-t</i>		
<i>bootstrap</i>	-0.05733 , -0.01127	-0.06282 , -0.00708
<i>paramet. boots.</i>	-0.05819 , -0.01168	-0.06337 , -0.00691

QUADRO 11 - VALOR DAS CORRECÇÕES DO ENV. E ACELER.

	z_0	a/EFR.	a/HALL
em β_1			
boot. conv.	-0.01244	-0.06160	-0.00905
boot. param.	-0.01009	-0.06160	-0.00776
em β_2			
boot. conv.	0.02382	0.05777	0.00178
boot. param.	0.00000	0.05777	0.00656
em β_3			
boot. conv.	0.11388	0.05883	0.00709
boot. param.	0.11094	0.05883	0.00751

Os I.C. têm amplitude inferior ao clássico, sendo o intervalo de menor amplitude o obtido com o método dos percentis, enquanto que o *bootstrap-t* e o BCa de Efron são os que têm amplitude mais próxima do intervalo clássico. Os intervalos deduzidos com o *bootstrap* paramétrico têm amplitude ligeiramente superior aos os outros I.C. *bootstrap*. No entanto, não existem grandes diferenças entre os intervalos determinados com o *bootstrap* convencional e o *bootstrap* paramétrico.

No quadro 11 encontram-se os valores das correcções do enviesamento e aceleração. Nota-se que as correcções actuam todas no mesmo sentido. A correcção de aceleração de Efron é bastante superior à de Peter Hall. A correcção do enviesamento em β_3 assume valores que se destacam nitidamente dos correspondentes aos outros coeficientes. Apesar do estimador *OMQ* não coincidir com o estimador da M.V. em pequenas amostras (conforme foi já referido), verifica-se para β_1 e β_2 uma proximidade entre a correcção de aceleração de Hall e a correcção do enviesamento.

Da comparação dos vários intervalos o facto mais importante a realçar consiste na verificação para os intervalos BCa e *bootstrap-t* da mesma tendência de desvio relativamente ao clássico (mesmo para os casos em que o método dos percentis denota uma tendência diferente), embora no *bootstrap-t* o desvio seja menos preponderante. Esta tendência caracteriza-se em β_1 e β_2 por um maior afastamento de um dos extremos do I.C. (em β_1 extremo superior e em β_2 extremo inferior) relativamente aos extremos do intervalo clássico. Por outro lado, a correcção de Peter Hall parece conduzir a intervalos mais "sensatos", situando-se mais próximos dos resultantes do *bootstrap-t*. Aponta-se ainda que para β_1 e β_3 a tendência de desvio dos percentis relativamente ao intervalo clássico é depois contrariada pelas correcções posteriores, donde os intervalos BCa apresentam um desvio de tendência oposta que é secundada pelo *bootstrap-t*. Este facto revela uma certa falta de precisão do método dos percentis, podendo produzir intervalos erróneos. A fraca *performance* deste método foi já detectada por vários autores em estudos de simulação que efectuaram.

FIG.4A - HISTOGRAMA DE $\hat{\beta}_1$ (BOOT)

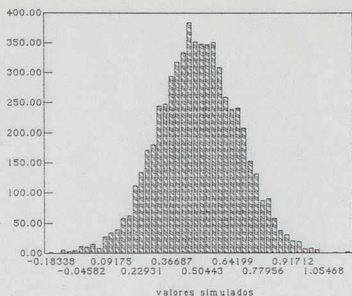


FIG.4B - HISTOGRAMA DE $\hat{\beta}_1$ (BOOT. PAR.)

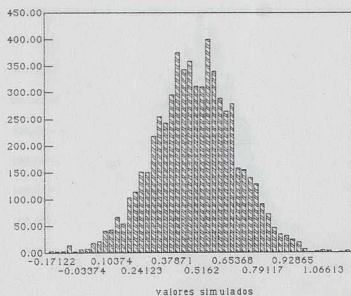


FIG.5A - HISTOGRAMA DE $\hat{\beta}_2$ (BOOT)

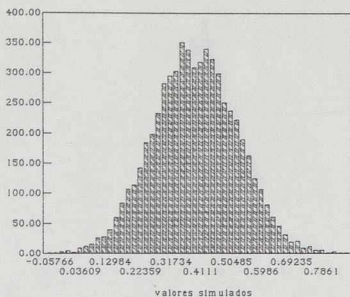


FIG.5B - HISTOGRAMA DE $\hat{\beta}_2$ (BOOT. PAR.)

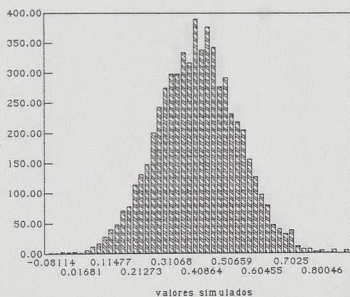


FIG.6A - HISTOGRAMA DE $\hat{\beta}_3$ (BOOT)

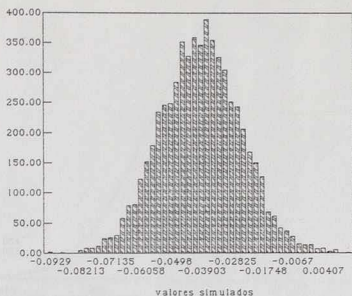


FIG.6B - HISTOGRAMA DE $\hat{\beta}_3$ (BOOT. PAR.)

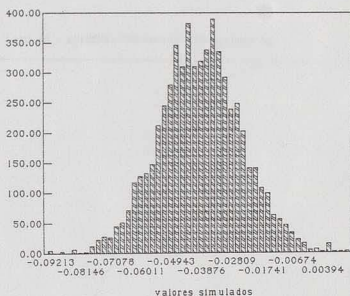
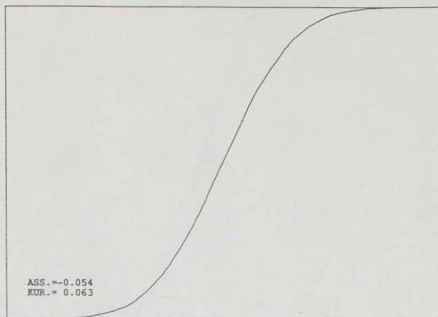
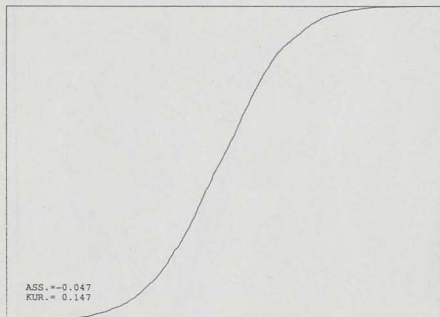


FIG.7A - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_1$

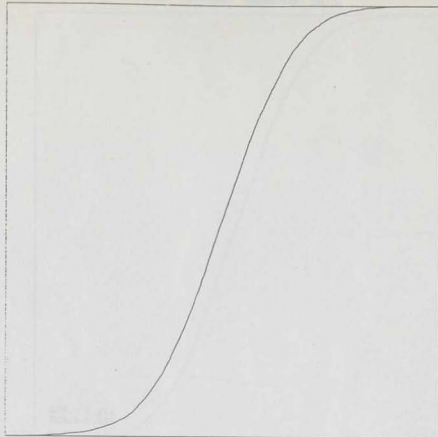


	min	max
X	-1.83377579E-01	1.19224286E+00
Y	0.00000000E+00	1.00000000E+00

FIG.7B - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_1$ (BOOT.PAR.)

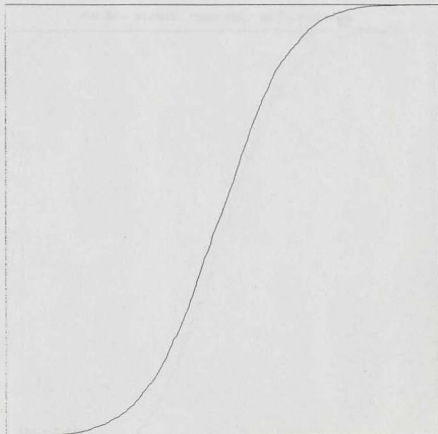


	min	max
X	-1.71224088E-01	1.20361924E+00
Y	0.00000000E+00	1.00000000E+00



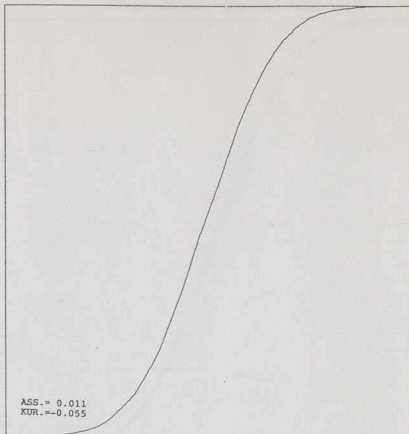
	min	max
X	-3.63338705E+00	3.92422633E+00
Y	0.00000000E+00	1.00000000E+00

FIG. 7D - DISTRIB. 'BOOTSTRAP' 'STUD.' DE $\hat{\beta}_2$ (BOOT.PAR.)



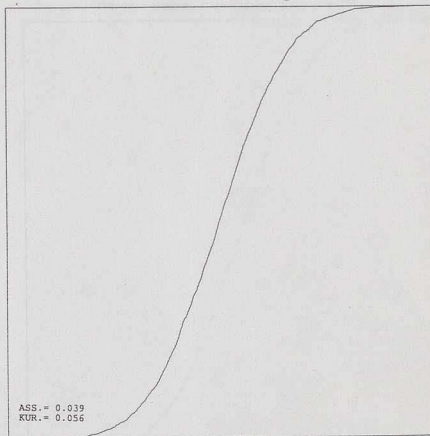
	min	max
X	-3.48590536E+00	3.89649931E+00
Y	0.00000000E+00	1.00000000E+00

FIG.8A - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_2$



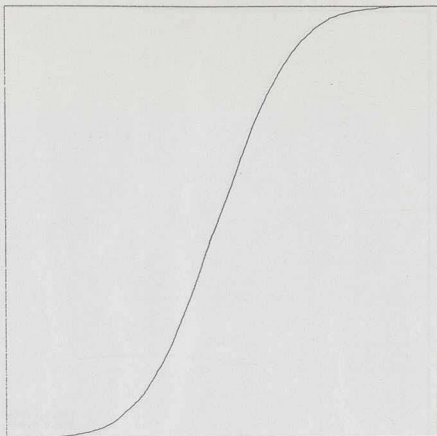
	min	max
X	-5.76637276E-02	8.79855871E-01
Y	0.00000000E+00	1.00000000E+00

FIG.8B - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_2$ (BOOT.PAR.)



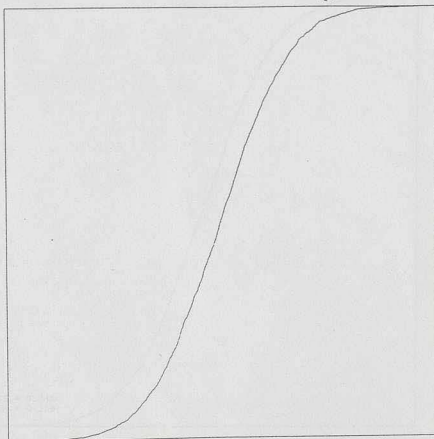
	min	max
X	-8.11408582E-02	8.98416162E-01
Y	0.00000000E+00	1.00000000E+00

FIG.8C - DISTRIB. 'BOOTSTRAP' 'STUD.' DE β_2



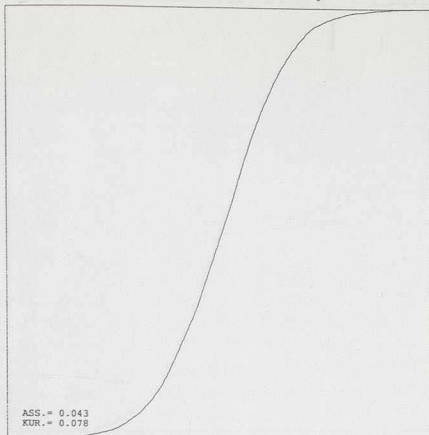
	min	max
X	-3.56103615E+00	3.64691489E+00
Y	0.00000000E+00	1.00000000E+00

FIG.8D - DISTRIB. 'BOOTSTRAP' 'STUD.' DE $\hat{\beta}_2$ (BOOT.PAR.)



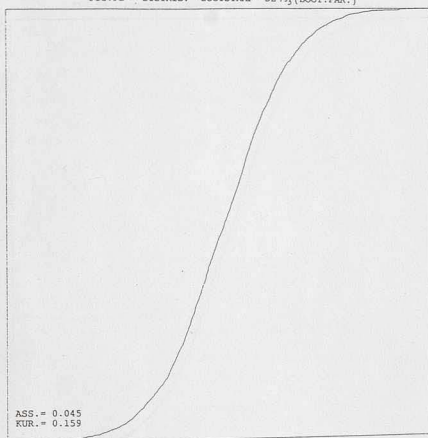
	min	max
X	-3.71500080E+00	3.76273603E+00
Y	0.00000000E+00	1.00000000E+00

FIG.9A - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_3$



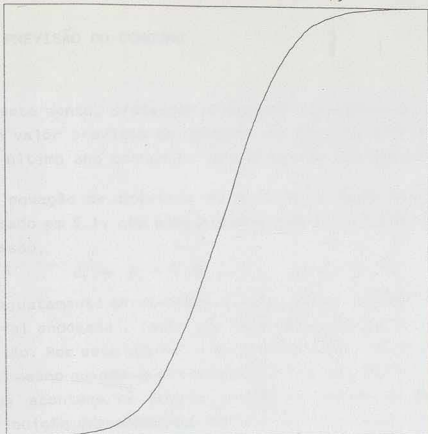
	min	max
X	-9.29010212E-02	1.48464832E-02
Y	0.00000000E+00	1.00000000E+00

FIG.9B - DISTRIB. 'BOOTSTRAP' DE $\hat{\beta}_3$ (BOOT.FAR.)



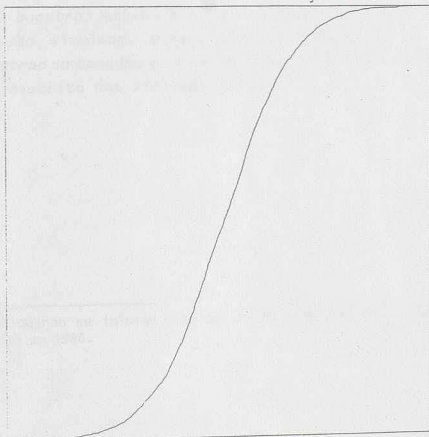
	min	max
X	-9.21304598E-02	1.46092596E-02
Y	0.00000000E+00	1.00000000E+00

FIG.9C - DISTRIB. 'BOOTSTRAP' 'STUD.' DE $\hat{\beta}_3$



	min	max
X	-3.97678399E+00	3.59381775E+00
Y	0.00000000E+00	1.00000000E+00

FIG.9D - DISTRIB. 'BOOTSTRAP' 'STUD.' DE $\hat{\beta}_3$ (BOOT.PAR.)



	min	max
X	-3.83405407E+00	3.49637312E+00
Y	0.00000000E+00	1.00000000E+00

5.2 - PREVISÃO DO CONSUMO

Neste ponto, pretende-se determinar intervalos de confiança para o valor previsto do Consumo nos anos de 1986 e 1987, visto que o último ano conhecido no período de estimação foi 1985.¹

A equação de previsão do Consumo é dada pelo ajustamento analisado em 5.1. que após algumas operações algébricas tem como expressão,

$$\hat{c}_t = \hat{\beta}_1 + \hat{\beta}_2 DY_{t-1} + c_{t-1} + \hat{\beta}_3 c_{t-2} + \hat{U}_t \quad (5.2)$$

Este ajustamento apresenta nos regressores valores desfasados da variável endógena, facto que terá repercussões perturbadoras na previsão. Por este motivo, o erro de previsão não tem distribuição Normal mesmo quando os erros têm distribuição Normal (ao contrário do que acontece no modelo clássico), sendo a sua verdadeira distribuição desconhecida. Por outro lado, a expressão analítica da variância do erro de previsão é bastante complexa tornando-se difícil de deduzir. Perante esta situação, os métodos clássicos de previsão tornam-se ineficazes na avaliação estatística dos valores previstos.

O *bootstrap* estima o erro de previsão através de um erro de previsão simulado, gerado pela diferença entre a estimativa *bootstrap* do Consumo e um valor actual simulado. Este procedimento será descrito nas próximas linhas.

¹ Quando se iniciou esta aplicação ainda não era conhecido o valor do Consumo em 1986.

5.1.1 - Previsão *bootstrap* para 1986

A previsão (pontual) para o Consumo em 1986 obtida com o *MOMQ* é, de acordo com (5.2), dada por,

$$\begin{aligned}\hat{c}_{86} &= \hat{\beta}_1 + \hat{\beta}_2 DY_{85} + c_{85} + \hat{\beta}_3 c_{84} \\ \hat{C}_{86} &= \exp \hat{c}_{86},\end{aligned}\quad (5.3)$$

onde os $\hat{\beta}_j$ ($j=1, \dots, 3$) correspondem às estimativas obtidas em 5.1 com o *MOMQ*.

Para proceder à previsão *bootstrap* é necessária a estimação prévia do modelo através do algoritmo *bootstrap* exposto no capítulo 4, cujos resultados foram analisados em (5.1.3). De seguida, há que efectuar o seguinte procedimento:

Em cada iteração b determinar,

1) A previsão *bootstrap*, \hat{C}_{86}^b , igual a,

$$\begin{aligned}c_{86}^b &= \hat{\beta}_1^b + \hat{\beta}_2^b DY_{85} + c_{85} + \hat{\beta}_3^b c_{84} \\ \hat{C}_{86}^b &= \exp c_{86}^b,\end{aligned}\quad (5.4)$$

em que os coeficientes $\hat{\beta}_j^b$ ($j=1, \dots, 3$) são os obtidos na estimação com a b reamostra (de acordo com o ponto 4 do algoritmo de 4.2);

2) O valor actual simulado, C_{86}^{*b} , dado por,

$$\begin{aligned}c_{86}^{*b} &= \hat{\beta}_1 + \hat{\beta}_2 DY_{85} + c_{85} + \hat{\beta}_3 c_{84} + \hat{U}_{86}^{*b} \\ C_{86}^{*b} &= \exp c_{86}^{*b},\end{aligned}\quad (5.5)$$

onde \hat{U}_{86}^{*b} é o resíduo resultante de uma extracção aleatória do conjunto dos 26 resíduos do ajustamento com o *MOMQ*. No *bootstrap* paramétrico \hat{U}_{86}^{*b} é gerado a partir da distribuição normal que determina a reamostragem.

3) O erro de previsão simulado, EPS_{86}^{*b} , de acordo com,

$$EPS_{86}^{*b} = C_{86}^{*b} - \hat{C}_{86}^b.$$

Repetir por B vezes esta iteração.

No final deste processo fica-se em poder de B valores simulados para o erro de previsão em 1986, $EPS'_{86}^1, \dots, EPS'_{86}^b, \dots, EPS'_{86}^z$. Aplicando a estes valores as técnicas *bootstrap* descritas em 2.1 é possível determinar as respectivas estimativas da média e desvio padrão. É ainda possível obter uma estimativa da distribuição amostral das previsões para o consumo através das frequências relativas acumuladas dos B valores do erro de previsão somados com a previsão do consumo dada pelo MOMQ.

5.2.2 - Previsão *bootstrap* para 1987

A previsão *bootstrap* para 1987 baseia-se num procedimento semelhante ao anterior, só que, neste caso, existe um problema adicional - é necessário o valor de DY_{86} , não conhecido no presente. Como tal há que substituí-lo por uma previsão. Com este fim, efectuou-se a modelização de $DY_t = y_t - y_{t-1}$, que consta no anexo 2. De acordo com os resultados patentes no anexo referido, DY_{86} é dado por,

$$DY_{86} = \hat{\phi}_1 + \hat{\phi}_2 DY_{85} \quad , \quad (5.6)$$

com $\hat{\phi}_1$ e $\hat{\phi}_2$ as estimativas determinadas em A2.2.

A previsão (pontual) dada pelo MOMQ para o Consumo em 1987 é igual a,

$$\hat{c}_{87} = \hat{\beta}_1 + \hat{\beta}_2 DY_{86} + \hat{c}_{86} + \hat{\beta}_3 c_{85}$$

$$\hat{C}_{87} = \exp \hat{c}_{87} \quad ,$$

com \hat{c}_{86} e DY_{86} dados respectivamente por (5.3) e (5.6).

Para obter a previsão *bootstrap* para 1987 há que proceder a B iterações, onde:

Em cada iteração b determinar,

1) A previsão *bootstrap*, \hat{C}_{87}^b , idêntica a,

$$\hat{C}_{87}^b = \hat{\beta}_1^b + \hat{\beta}_2^b DY_{86}^b + \hat{c}_{86}^b + \hat{\beta}_3^b c_{85}$$

$$DY_{86}^b = \hat{\phi}_1^b + \hat{\phi}_2^b DY_{85}$$

$$\hat{C}_{87}^b = \exp \hat{c}_{87}^b .$$

Esta equação necessita de alguns esclarecimentos adicionais. É pressuposto que se efectua a DY_i uma estimação *bootstrap* nos mesmos moldes que a implementada em 5.1.3 a DC_i . O valor de DY_{86}^b decorre do ajustamento a DY_i quando se utilizam os dados da reamostragem b. Dado que foi aceite a hipótese que as variáveis residuais da regressão de DY_i têm distribuição Normal (vide Anexo 2) é possível aplicar o *bootstrap* paramétrico. Assim, na previsão com o *bootstrap* paramétrico, as reamostragens (na regressão de DY_i são definidas a partir da Normal de média zero e desvio padrão fixado no valor da estimativa dada pelo MOMQ para o desvio padrão de ϵ_i (sendo ϵ_i a variável residual que modeliza DY_i). Os coeficientes $\hat{\beta}_j^b$ são os mesmos que anteriormente, sendo \hat{c}_{86}^b dado por (5.5).

2) O valor actual simulado, C_{87}^{*b} , sendo,

$$C_{87}^{*b} = \hat{\beta}_1 + \hat{\beta}_2 DY_{86}^{*b} + c_{86}^{*b} + \hat{\beta}_3 c_{85} + \hat{U}_{87}^{*b}$$

$$DY_{86}^{*b} = \hat{\phi}_1 + \hat{\phi}_2 DY_{85} + \hat{\epsilon}_{86}^{*b}$$

$$C_{86}^{*b} = \exp c_{87}^{*b} ,$$

onde c_{86}^{*b} é determinado por (5.5), \hat{U}_{87}^{*b} resulta de uma extracção aleatória do conjunto dos 26 resíduos da regressão do MOMQ de DC_i e $\hat{\epsilon}_{86}^{*b}$ é um resíduo extraído aleatoriamente do conjunto dos resíduos do ajustamento a DY_i explicitado no anexo 2. No *bootstrap* paramétrico \hat{U}_{87}^{*b} e $\hat{\epsilon}_{86}^{*b}$ são gerados a partir das respectivas distribuições Normal que definem a reamostragem de cada um dos ajustamentos.

3) O erro de previsão simulado, EPS_{87}^{*b} , igual a,

$$EPS_{87}^{*b} = C_{87}^{*b} - \hat{C}_{87}^b.$$

Repetir B vezes esta iteração.

Na posse dos B valores simulados do erro de previsão obtêm-se, tal como para 1986, as estimativas *bootstrap* da média e do desvio padrão desta variável, sendo também possível uma aproximação da distribuição empírica das previsões. Esta corresponde às frequências acumuladas dos B valores para ao erro de previsão somados com a previsão do consumo dada pelo *MOMQ*.

Implementando os procedimentos descritos obtiveram-se as estimativas *bootstrap* da média e desvio padrão do erro de previsão em 1986 e 1987, patentes no Quadro 12, enquanto que a previsão do Consumo dada pelo *MOMQ* para 1986 e 1987 foi respectivamente igual a 446 258.7 10³ contos e 474 994.8 10³ contos.

QUADRO 12 - ESTIM. *BOOTSTRAP* DA MÉDIA E D.P. DO ERRO DE PREV.

	1986		1987	
	MÉDIA	D.P.	MÉDIA	D.P.
<i>bootstrap</i>	190.0	10 607.6	14.3	17 443.0
<i>boot.par.</i>	779.2	15 992.1	1 197.8	25 104.9

Nota-se uma diferença relevante entre os resultados dados pelo *bootstrap* convencional e o *bootstrap* paramétrico, sendo estes últimos bastante superiores. O erro de previsão é positivo mostrando que as previsões do *MOMQ* têm tendência a subavaliar os verdadeiros valores. No entanto, é possível concluir que o erro de previsão estimado não é muito significativo. É estranho o erro de previsão do *bootstrap* convencional descer em 1987, pois a incerteza neste ano é maior visto que se está mais longe do período de estimação, havendo mesmo a necessidade de entrar com

valores previstos nas variáveis explicativas. Desta forma, não é de estranhar que em 1987 o desvio padrão do erro de previsão assuma valores mais elevados.

Os I.C. para as previsões foram deduzidos a partir das distribuições *bootstrap* empíricas determinadas para o Consumo em 1986 e 1987. Não foi possível determinar o *bootstrap-t* porque não existe uma estimativa para a variância do erro de previsão distinta da *bootstrap*.

Na obtenção da constante de aceleração pela via de Efron pensou-se em considerar

$$U_i = (N-1) [EPS(.) - EPS(i)] ,$$

$$EPS(.) = \frac{1}{N} \sum_{i=1}^N EPS(i) ,$$

onde $EPS(i)$ foi calculado através do algoritmo de previsão descrito substituindo em 2) os coeficientes β_j^b pelos $\beta_{j,i}$ ($j=1,2,3$), obtidos com o ajustamento sem as observações do período i ; para 1987 há que substituir também em 2) os ϕ_j^b pelos $\phi_{j,i}$ ($j=1,2$) dados pela regressão de DY_i sem as observações do período i .

O que mais surpreendeu nestes resultados foi a tendência oposta entre as correcções do enviesamento e aceleração de Efron. É curioso que, neste caso, as correcções quase que se anulam aproximando o intervalo BCa do dado pelos percentis. No *bootstrap* convencional esta oposição não se verifica com a correcção da aceleração de Hall, apresentando mesmo um valor aproximado ao de z_0 . No entanto, para o *bootstrap* paramétrico a constante de aceleração de Hall diminui bastante o seu valor absoluto,

tornado-se menos efectiva. A opposição entre as correcções do enviesamento e aceleração não tem explicação aparente a não ser que reflecta uma certa falta de precisão na estimação de a .¹

Nota-se uma certa diferença nos resultados obtidos com o *bootstrap* paramétrico em relação aos do convencional. Para além da já referida alteração da constante de Hall, os I.C. dados pela variante paramétrica têm maior amplitude sendo centrados num valor ligeiramente superior, embora esta amplitude seja razoável (cerca de 20% do extremo inferior).

Os histogramas e as distribuições empíricas das previsões do Consumo para 1986 e 1987 encontram-se nas figuras 10, 11, 12 e 13 respectivamente. É de notar que a distribuição em 1987 apresenta maior regularidade do que em 1986.

1 A determinação de a em problemas de previsão deste tipo é um assunto não abordado pela literatura consultada para este trabalho, quer em termos teóricos quer empíricos. É pois natural que a generalização efectuada para esta aplicação não gose das propriedades teóricas da constante de aceleração analisadas no capítulo 3.

QUADRO 13 - INTERV. DE CONFIANÇA PARA A PREV. DO CONSUMO 1986

	90%	95%
<i>método percen.</i>		
<i>bootstrap</i>	446 942.0 , 481 621.8	441 745.3 , 483 801.5
<i>paramet. boots.</i>	440 908.9 , 493 270.4	436 200.0 , 498 751.1
<i>correc. envies.</i>		
<i>bootstrap</i>	444 739.9 , 480 707.9	440 635.9 , 482 901.2
<i>paramet. boots.</i>	439 723.9 , 492 403.3	435 157.6 , 497 569.6
<i>c.env/ac.-EFRON</i>		
<i>bootstrap</i>	446 672.7 , 481 531.3	442 200.9 , 484 054.0
<i>paramet. boots.</i>	441 338.8 , 494 042.1	436 938.5 , 500 053.3
<i>c.env./ac.-HALL</i>		
<i>bootstrap</i>	441 221.3 , 479 409.9	437 701.3 , 481 262.9
<i>paramet. boots.</i>	439 895.7 , 492 483.3	435 484.3 , 497 831.6

QUADRO 14 - INTERV. DE CONFIANÇA PARA A PREV. DO CONSUMO 1987

	90%	95%
<i>método percen.</i>		
<i>bootstrap</i>	445 436.2 , 502 521.8	439 175.4 , 507 054.5
<i>paramet. boots.</i>	435 940.0 , 517 556.1	427 553.2 , 526 169.6
<i>correc. envies.</i>		
<i>bootstrap</i>	444 394.9 , 502 633.4	437 789.1 , 506 419.3
<i>paramet. boots.</i>	434 794.3 , 516 095.5	426 077.0 , 525 003.8
<i>c.env/ac.-EFRON</i>		
<i>bootstrap</i>	445 156.1 , 502 425.7	439 352.3 , 507 232.8
<i>paramet. boots.</i>	435 692.9 , 517 373.1	427 715.1 , 526 448.9
<i>c.env./ac.-HALL</i>		
<i>bootstrap</i>	442 650.5 , 500 937.2	435 512.7 , 504 755.0
<i>paramet. boots.</i>	435 442.1 , 516 778.8	427 015.3 , 525 709.6

QUADRO 15 - VALOR DAS CORRECÇÕES DO ENV. E ACELER.

	z_0	a/EFR.	a/HALL
em 1986			
<i>bootstrap</i>	-0.0589	0.0371	-0.0769
<i>boot. param.</i>	-0.0301	0.0371	0.0044
em 1987			
<i>bootstrap</i>	-0.0267	0.0158	-0.0310
<i>boot. param.</i>	-0.0255	0.0158	0.0102

A conjuntura económica em 1986 e 1987 induziu uma política económica expansionista com repercussões no crescimento do Consumo. De acordo com os relatórios do Banco de Portugal a taxa do crescimento real do Consumo foi em 1986 de 6.8% e em 1987 de 7%, valores bastante elevados. Apicando estas taxas de crescimento ao valor conhecido das Despesas de Consumo de 1985, obtêm-se os valores para as Despesas de Consumo em 1986 e 1987, respectivamente de 492 062.0 e 525 522.2. As Despesas de Consumo em 1986 caí dentro do I.C. a 90% dado com o *bootstrap* paramétrico enquanto que as Despesas de Consumo em 1987 estão incluídas no I.C. a 95% também obtido com a versão paramétrica. Os extremos superiores dos I.C. deduzidos com o *bootstrap* convencional, apesar de não conterem estes valores, estes encontram-se muito próximos dos extremos superiores daqueles.

Estes resultados podem ser considerados bastante satisfatórios não só porque a equação de previsão não era a ideal (apresentando um coeficiente de determinação não elevado), como também os valores verificados foram bastante elevados tendo sido influenciados por factos conjunturais não incorporados na informação do período de estimação. Estes valores são nitidamente superiores às previsões do MOMQ. Apesar destes factos, os I.C. *bootstrap* (versão paramétrica) contêm os valores verificados.

FIG. 10A - HISTOGRAMA DA PREVISÃO/1986

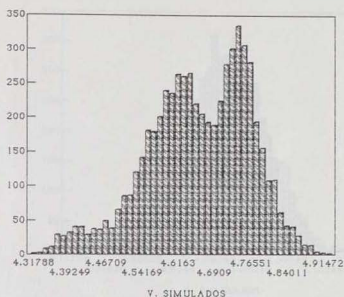


FIG.10B - HIST. PREVISÃO/1986 (B.PAR)

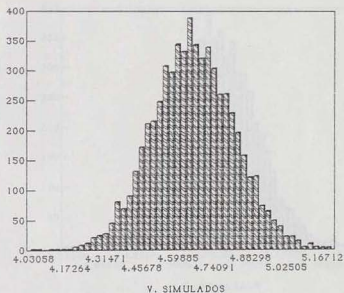


FIG. 11A - HISTOGRAMA DA PREVISÃO/1987

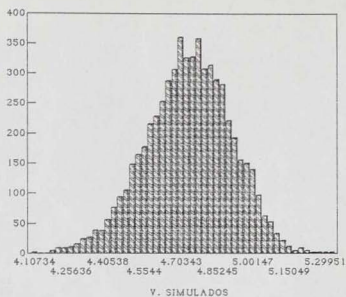


FIG.11B - HISTOG. PREVISÃO/1987 (B.PAR)

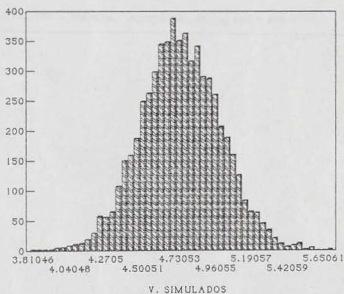
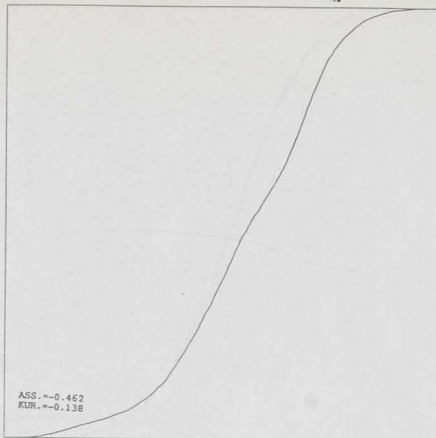
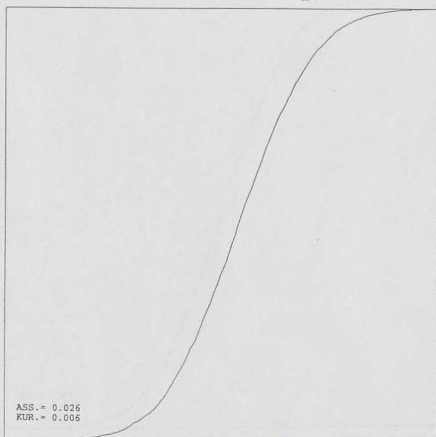


FIG.12A - DISTRIB. 'BOOTSTRAP' DE \hat{C}_{16}



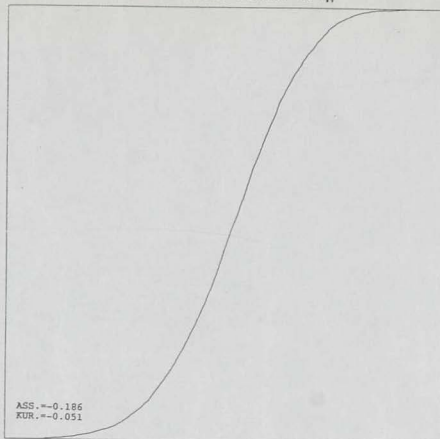
	min	max
X	4.31788060E+05	4.93958660E+05
Y	0.00000000E+00	1.00000000E+00

FIG.12B - DISTRIB. 'BOOTSTRAP' DE \hat{C}_{16} (BOOT.PAR.)



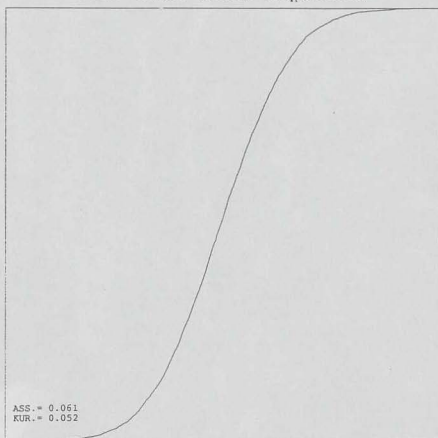
	min	max
X	4.03057630E+05	5.21447060E+05
Y	0.00000000E+00	1.00000000E+00

FIG.13A - DISTRIB. 'BOOTSTRAP' DE \hat{C}_{11}



	min	max
X	4.10733940E+05	5.34918750E+05
Y	0.00000000E+00	1.00000000E+00

FIG.13B - DISTRIB. 'BOOTSTRAP' DE \hat{C}_{11} (BOOT.PAR.)



	min	max
X	3.81045880E+05	5.72727940E+05
Y	0.00000000E+00	1.00000000E+00

A evolução recente da estatística induzida pelo desenvolvimento da informática, e que se caracteriza pela predominância cada vez maior da utilização de técnicas de simulação com computação intensiva, conduziu à emergência de novos métodos como o *bootstrap*.

O *bootstrap*, ao substituir a modelização e formalização teóricas pelo recurso à computação intensiva, aplica-se a um conjunto vasto de situações, permitindo a determinação de forma simples e automática de soluções (geralmente numéricas) em problemas de resolução demasiado complexa ou mesmo insolúveis no quadro da teoria clássica. Por outro lado, além do seu carácter operacional possui ainda propriedades teóricas assinaláveis. É assintoticamente válido e, em muitos problemas, conduz a aproximações de melhor qualidade do que as aproximações clássicas com base no Teorema do Limite Central. Com estas características, o *bootstrap* não pode deixar de constituir actualmente um instrumento indispensável para os estatísticos.

Com este trabalho pretendeu mostrar-se como pode ser vantajosa a aplicação do *bootstrap* em modelos econométricos. Por um lado, quando não existem resultados estabelecidos para amostras finitas, pode constituir uma alternativa mais eficiente à aplicação directa dos limites assintóticos conhecidos. Por outro lado, pode proporcionar uma solução para problemas não resolvidos com os métodos tradicionais.

Na primeira situação, enquadra-se a determinação de I.C. *bootstrap* para os parâmetros da função consumo como alternativa aos baseados na distribuição assintótica *t* de *student*. Os resultados obtidos mostram que os I.C. clássicos centrados na estimativa dada com o MOMQ podem ser algo erróneos. Verificou-se,

nós I.C. BC, BC_a e *bootstrap-t*, coerência nos resultados, caracterizando-se estes pelo afastamento nítido de um dos extremos relativamente ao intervalo clássico. Neste caso, pensa-se que os custos computacionais adicionais exigidos pelo *bootstrap* são compensatórios, por permitirem complementar e precisar a informação decorrente dos resultados assintóticos clássicos.

Na segunda situação, encontra-se a determinação da estimativa do enviesamento do estimador *OMQ*, e a avaliação estatística das previsões com a estimação da média e desvio padrão do erro de previsão e ainda a construção de I.C. para as Despesas de Consumo em 1986 e 1987.

Sabe-se que os estimadores *bootstrap* do enviesamento e da variância de β , são consistentes conhecendo-se a respectiva taxa de convergência, permitindo uma certa segurança na interpretação dos resultados obtidos na aplicação empírica efectuada. Assim é possível concluir que o enviesamento do estimador *OMQ* na determinação das estimativas para os parâmetros da função consumo é negligenciável, verificando-se uma grande semelhança entre as estimativas da variância dos β , dadas pelo *MOMQ* e as estimativas *bootstrap*.

Na previsão, as soluções *bootstrap* obtidas mostram-se credíveis, sendo bastante satisfatórias. Estas, dependem do preditor utilizado, pois é de notar que o *bootstrap* não constitui um método de previsão, sendo antes um método que permite avaliar a variabilidade estatística do preditor. Apesar do valor verificado para as Despesas de Consumo se distanciar da previsão dada pelo *MOMQ*, os I.C. *bootstrap* (obtidos com o *bootstrap* paramétrico) contêm este valor.

As previsões poderão ser melhoradas com a introdução adicional de algumas variáveis explicativas na função preditiva. Esta estratégia não foi prosseguida por se repercutir numa complexidade demasiado elevada da computação na implementação do *bootstrap*, extravasando os objectivos deste trabalho.

Os resultados do *bootstrap* paramétrico não diferem significativamente dos obtidos com o *bootstrap* convencional, denotando em regra uma maior variabilidade. Em termos formais o *bootstrap* paramétrico parece ser mais apropriado para a aplicação em causa visto que, a reamostragem a partir de uma Normal respeita mais fielmente a estrutura probabilística identificada no problema. É de notar que na previsão o *bootstrap* paramétrico apresentou uma *performance* superior.

Relativamente à construção de I.C. aproximados para os parâmetros do modelo destaca-se o *bootstrap-t* por ter uma calculatória mais simples, dado que não necessita da determinação de correcções analíticas tais como a correcção do enviesamento e a da aceleração, apresentando resultados semelhantes aos do BCa. Em particular neste problema, onde a distribuição *t* de *student* é válida assintoticamente, é de crer, intuitivamente, que a *studentização* em amostras finitas conduza a bons resultados.

As propriedades dos intervalos BCa dependem da determinação da constante de aceleração. Na aplicação empírica foram utilizadas duas estimativas. A de Efron com base na função de influência empírica de β_j e a de Peter Hall que recorre ao coeficiente de assimetria empírico da distribuição *bootstrap* de $\sqrt{n}(\hat{\beta}_j^* - \beta_j)$. Embora esta última tenha sido deduzida para os problemas específicos em que a média e a variância da estatística são funções suaves da média da amostra, a sua aplicação no problema de regressão em causa surtiu bons resultados, conduzindo a intervalos mais "sensatos" (na medida em que se aproximam mais do *bootstrap-t* e não estão tão desviados do clássico) do que os obtidos com a constante de aceleração de Efron. É de referir que, com o *bootstrap* paramétrico, não foi utilizada a estimativa indicada por Efron em ambiente paramétrico dada por (3.42) visto a sua determinação ser demasiado complexa. Na previsão foi mais relevante a superioridade da correcção de Hall na obtenção de bons I.C..

A aplicação de métodos como a calibragem de Loh e a pré-pivotação de Beran poderá melhorar a cobertura real dos I.C. aproximados. Contudo, esta aplicação envolve custos de computação bastante elevados, particularmente em problemas de regressão. No entanto, surgiram muito recentemente técnicas de reamostragem que permitem uma substancial diminuição dos cálculos a efectuar, como a *Importance sampling* e *Antithetic resampling*¹.

A análise valorativa da qualidade dos intervalos *bootstrap* que transparece nestas considerações é essencialmente heurística, dado que nos problemas de regressão não estão devidamente estudadas as propriedades da aproximação *bootstrap* aos intervalos correctos. Estas propriedades podem ser deduzidas com a aplicação de Expansões de Edgeworth e Cornish Fisher às distribuições dos estimadores dos parâmetros e respectivos quantis.

Os temas que constam neste dois últimos parágrafos constituem linhas de investigação que se pensa desenvolver em trabalhos futuros.

¹ Sobre estas técnicas veja-se:
Jonhs, M.V.(1988). Importance Sampling for Bootstrap Confidence Intervals. *J.A.S.A.*, 403,83, 709-714.
Hinkley, D.V.; Shi, S.(1989). Importance Sampling and the Nested Bootstrap. *Biometrika*, 76, 3, 435-446.
Hall, P.(1989). Antithetic Resampling for the Bootstrap. *Biometrika*, 76, 4, 713-724.

ANEXO 1 - TESTES DE ESPECIFICAÇÃO DO MODELO

A1.1 - DETECÇÃO DE AUTOCORRELAÇÃO

A1.1.1 - Estatística de Durbin para modelos dinâmicos

A estatística de Durbin de 1970 para modelos com variável endógena desfasada (D70) permite testar a hipótese de inexistência de autocorrelação nas variáveis residuais do modelo contra a alternativa de estas verificarem um processo AR(1) ou MA(1). Esta estatística tem distribuição assintótica $t_{(N-k-2)}$. É de referir que a potência do teste dado por D70 é nitidamente afectada em pequenas amostras.

A1.1.2 - Teste de Godfrey

Este é um teste deduzido através da técnica dos multiplicadores de Lagrange e abrange modelos com variável endógena desfasada. Godfrey(1978) mostra que a estatística que define o teste é assintoticamente equivalente a,

$$GOD = N(r_1^2 + r_2^2 + \dots + r_p^2) ,$$

com r_i uma estimativa do i -ésimo coeficiente de autocorrelação dos erros, calculada a partir dos resíduos da regressão com o MOMQ. Esta estatística tem distribuição assintótica $\chi^2_{(p)}$.

O teste de Godfrey permite ensaiar a hipótese de que os U_i têm uma distribuição $\mathcal{N}(0, \sigma^2)$ ($i=1, \dots, N$) contra a alternativa de serem um processo AR(p) ou MA(p).

A1.2 - DETECÇÃO DE HETEROESCEDASTICIDADE

A1.2.1 - TESTE DE ENGLE

Este teste ensaia a hipótese de homoescedasticidade dos erros contra a alternativa de estes constituírem um processo ARCH (heteroescedasticidade condicional autoregressiva). É um teste tipo multiplicadores de Lagrange.

O processo ARCH caracteriza-se por ter média nula, ser não correlacionado no tempo, com variâncias condicionadas pelos períodos anteriores não constantes e variâncias não condicionadas constantes. O modelo ARCH de ordem p traduz-se por,

$$U_t | \Psi_{t-1} \sim \mathcal{N}(0, h_t) ,$$

$$h_t = h(U_{t-1}, \dots, U_{t-p}, \underline{\alpha}) ,$$

com Ψ_{t-1} o conjunto de informação disponível no período t e $\underline{\alpha}$ um vector de parâmetros desconhecidos. O teste de Engle pode ser aplicado em modelos com variável endógena desfasada. Neste caso, em presença de erros ARCH as estimativas da variância dadas pelo MOMQ não são consistentes, enquanto que o estimador da M.V. é não linear e é mais eficiente que o MOMQ.

A estatística do teste é assintoticamente equivalente a NR^2 e tem distribuição limite $X^2_{(p)}$, sendo R^2 o coeficiente de correlação múltipla da regressão de \hat{U}_t^2 sobre $\hat{U}_{t-1}^2, \dots, \hat{U}_{t-p}^2$ mais uma constante. Este teste é localmente assintoticamente mais potente.

A1.2.2 - Teste de White

O teste de White tem por base a comparação entre o estimador da matriz de covariâncias de β dado pelo MOMQ e o estimador deduzido pelo autor que é consistente com heteroscedasticidade. Com homoscedasticidade dos erros ambos os estimadores terão que produzir estimativas muito próximas.

Este teste permite ainda inferir sobre a especificação do modelo pois a hipótese nula estipula, conjuntamente com a homoscedasticidade dos erros, que o valor esperado da variável endógena ($X\beta$) está correctamente definido. O teste de White tem a vantagem de não exigir nenhuma especificação particular para o tipo de heteroscedasticidade da hipótese alternativa.

Assumindo adicionalmente que $E\{U_i^4\} = \mu_4$, simplifica o processamento do teste. Neste caso, considerando a regressão em causa no capítulo 5, a implementação do teste passa pela realização da regressão artificial de U_i^2 numa constante e nas variáveis $X_{i2}, X_{i3}, X_{i2}^2, X_{i2}X_{i3}$, e X_{i3}^2 , onde X_{i2}, X_{i3} , são as variáveis explicativas da equação mencionada.

A1.2.3 - Teste de Breush e Pagan

Este é mais um teste de multiplicadores de Lagrange, obtido através da maximização da verosimilhança sujeita às restrições dos parâmetros decorrentes da hipótese nula.

O teste detecta a existência de heteroscedasticidade do tipo $\sigma^2 = h(Z_i' \alpha)$ onde Z_i é um vector (px1) em que o primeiro elemento é igual à unidade e α é um vector (px1) de parâmetros desconhecidos, funcionalmente não relacionados com os β , ($j=1, \dots, k$). A hipótese nula fixa $\alpha_1 = \alpha_2 = \dots = 0$ traduzindo a homoscedasticidade dos erros. Para aplicar o teste há que escolher a forma funcional de Z_i . De seguida efectua-se uma regressão

artificial em que a variável endógena é $(\sqrt{N}U_t^2)/\sqrt{SSR}$ com SSR a soma do quadrado dos resíduos, e os regressores são dados por \underline{Z}_t . A estatística do teste é igual a metade da soma dos quadrados explicada na regressão mencionada e tem distribuição assintótica $\chi^2_{(p-1)}$. A potência do teste é afectada em pequenas amostras (embora nada indique que seja pior do que a da generalidade dos testes sobre heteroscedasticidade), mostrando um enviesamento da ordem $O(N^{-1/2})$.

Neste trabalho escolheu-se duas especificações para \underline{Z}_t , \underline{Z}_t^1 e \underline{Z}_t^2 , com $\underline{Z}_t^1 = (1, X_1, X_2)$ e $\underline{Z}_t^2 = (1, \ln X_1, \ln X_2)$, com X_1 e X_2 as variáveis explicativas da equação tratada no capítulo 5.

A1.3 - TESTES À NORMALIDADE DOS ERROS

A1.3.1 - Análise da assimetria e do kurtosis

Para averiguar a normalidade dos erros foi analisado o valor dos coeficientes empíricos de assimetria, $\sqrt{b_1}$, e kurtosis, b_2 . Os respectivos valores foram confrontados com os valores críticos apresentados por White e MacDonald.¹

Por outro lado, também se avaliou a estatística $(N/26)[\sqrt{b_1}]^2 + (N/24)b_2^2$ que, como é sabido, sob a hipótese de normalidade tem distribuição assintótica $\chi^2_{(2)}$.

1 White e MacDonald(1980) p.20

A1.3.2 - Teste de Shapiro Wilk

A estatística do teste é calculada com base nos resíduos obtidos com o MOMQ, sendo igual a,

$$SW = \frac{\left(\sum_{i=1}^N \alpha_{iN} \hat{U}_{iN} \right)^2}{\sum_{i=1}^N \hat{U}_i^2},$$

com \hat{U}_{iN} a i-ésima estatística de ordem dos resíduos, e α_{iN} coeficientes tabulados por Shapiro e Wilk(1965). Os valores críticos assumidos por SW sob a hipótese de normalidade foram retirados do mesmo texto (p.605).

Shapiro e Wilk(1965) concluem que o teste reage a diversos aspectos de não normalidade mesmo em pequenas amostras com $N=20$, sendo especialmente sensível à assimetria e a abas longas. Esta estatística geralmente tem potência superior à de $\sqrt{b_1}$ ou b_2 .

A1.3.3 - Teste de Shapiro Francia

O teste de Shapiro Francia representa uma variante do teste anterior, sendo prosseguido através da determinação da estatística,

$$SF = \frac{\left(\sum_{i=1}^N [m_{iN} / (m_N^t m_N)^{1/2}] \hat{U}_{iN} \right)^2}{\sum_{i=1}^N \hat{U}_i^2},$$

onde \hat{U}_{iN} tem o mesmo significado que em SW, e m_N é um vector ($N \times 1$) em que o elemento m_{iN} ($i=1, \dots, N$) traduz o valor esperado da i-ésima estatística de ordem de uma amostra retirada de uma

Normal estandardizada n-dimensional. O valor destes coeficientes foi retirado de Harter.¹ Os pontos críticos de SF são bem aproximados pelos de SW em problemas com $N \leq 50$.

A sensibilidade de SF é similar à de SW, denotando apenas pequenas diferenças que dependem das características da distribuição alternativa. Ambos os testes são assintoticamente equivalentes.

A1.3.4 - Teste de Kiefer e Salmon

Este teste permite ensaiar a hipótese de se verificar simultaneamente $\mu_3 = 0$ e $\mu_4 = 0$. O teste é deduzido a partir da função de verosimilhança para um conjunto de N observações da variável residual, aplicando uma expansão de Edgeworth à distribuição desta variável. A estatística obtida é igual a,

$$KS = \frac{N}{6}(\hat{\mu}_3 - 3\hat{\mu}_1)^2 + \frac{N}{24}(\hat{\mu}_4 - 6\hat{\mu}_2 + 3)^2$$

com $\hat{\mu}_i$ o i-ésimo momento central empírico dos resíduos. A distribuição assintótica de KS é uma $\chi^2_{(2)}$. A hipótese nula pode ser particionada, ensaiando apenas $\mu_3 = 0$, através da primeira parcela de KS que tem distribuição assintótica $\chi^2_{(1)}$, ou somente $\mu_4 = 0$, com a segunda parcela de KS distribuída da mesma forma. Deste modo, o teste de Kiefer e Salmon tem a vantagem de permitir uma melhor compreensão da natureza do desvio da normalidade.

¹ Harter, L. (1965). Expected values of Normal order statistics. *Biometrika*, **48**, 151-165.

A1.3.5 - Teste de Jarque e Bera

O teste de Jarque e Bera permite ensaiar a hipótese conjunta de normalidade, homoescedasticidade e não correlação temporal. É mais um teste de multiplicadores de Lagrange. A estatística que define o teste é igual a,

$$LM_{NHI} = N \left[\frac{(\sqrt{b_1})^2}{6} + \frac{b_2^2}{24} \right] + BP1 + N(\underline{\hat{r}}^t \underline{\hat{r}}) ,$$

com BP1 a estatística de Breush e Pagan para \underline{Z}_i^1 de acordo com o referido em A1.2.3 e $\underline{\hat{r}} = [\hat{r}_i]$ o vector (px1) em que \hat{r}_i é a autocorrelação empírica de ordem i dos resíduos. A primeira parcela de LM_{NHI} , LM_N , corresponde ao teste de normalidade apresentado em A1.3.1 com distribuição assintótica $\chi_{(2)}^2$, a segunda parcela LM_H , traduz o teste de Breush e Pagan com distribuição assintótica $\chi_{(q)}^2$, e a última LM_I é igual à estatística de Godfrey dada em A1.1.2 que tem distribuição limite $\chi_{(p)}^2$. A estatística LM_{NHI} distribui-se assintoticamente de acordo com $\chi_{(2+q+p)}^2$. O teste tem a vantagem de ser aditivo o que permite outras especificações para a hipótese nula. Por exemplo, admitindo a normalidade pode testar-se conjuntamente a heteroscedasticidade e autocorrelação através de $LM_H + LM_I$ que segue uma $\chi_{(q+p)}^2$.

Este teste é assintoticamente equivalente ao teste com base na razão de verosimilhanças sendo por isso localmente mais potente em grandes amostras.

A1.4 - DETECÇÃO DE OBSERVAÇÕES INFLUENTES

A1.4.1 - Análise dos w_i

Seja

$$w_i = X_i' (X'X)^{-1} X_i.$$

O valor de w_i traduz a distância das observações do período i em relação à média dos dados. Quanto maior o valor de w_i , maior a preponderância da observação i da variável endógena no ajustamento, traduzindo também um maior afastamento em relação à média verificado pelas observações das variáveis explicativas neste período. Assim, valores elevados de w_i reflectem observações influentes ou *outliers*. Num modelo equilibrado, em que as observações têm igual influência, os valores de w_i ($i=1, \dots, N$) serão iguais a k/N . Se algum w_i for superior a $2k/N$, ou com mais certeza a $3k/N$, as observações i podem ser consideradas influentes. Huber¹ aponta ainda que valores para $w_i \leq 0.2$ são valores seguros, $0.2 \leq w_i \leq 0.5$ apresentam algum risco, e valores da ordem $w_i > 0.5$ devem ser evitados.

A1.4.2 - Análise do índice $DFITS_i$

Este índice traduz a influência nos valores estimados do conjunto das observações do período i , sendo igual a,

$$DFITS_i = \frac{\hat{Y}_i - \hat{Y}(i)_i}{\hat{\sigma}(i)\sqrt{w_i}} = \frac{w_i \hat{U}_i}{\hat{\sigma}(i)(1-w_i)\sqrt{w_i}},$$

onde $\hat{Y}(i)_i$ denota o valor ajustado da variável endógena na regressão sem o conjunto das observações do período i e $\hat{\sigma}(i)$ a

¹ Veja-se o capítulo 7 de Huber, P.J.(1981); *Robust statistics*; Wiley; New York.

estimativa para σ obtida nesta regressão. Em cada período i calcula-se um valor para $DFITS_i$. Quanto maior for este valor, maior a influência das observações i .

A1.4.3 - Função de influência empírica para β_j

A função de influência empírica da observação i é igual a,

$$IF(i) = \hat{\beta}_j(i) - \hat{\beta}_j$$

com $\hat{\beta}_j(i)$ a estimativa obtida com o *MOMQ* na regressão sem as observações do período i . $IF(i)$ mede a variação na estimativa de β_j quando se retiram as observações do período i .

A1.4.4 - Teste $OUT(i)$

A estatística $OUT(i)$ constrói-se com base no resíduo *studentizado* que é igual a $\hat{u}_i = \hat{U}_i / [\hat{\sigma}(\sqrt{1-w_i})]$, sendo dada por,

$$OUT(i) = \hat{u}_i \left(\frac{N-k-1}{N-k-\hat{u}_i^2} \right)^{1/2}$$

Este teste permite ensaiar a hipótese de não existência de observações influentes. Os valores correspondentes aos diversos níveis de significância encontram-se tabelados em Weisberg(1980) pp.266 e 267.

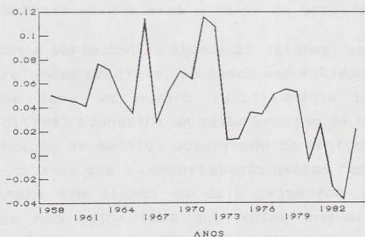
ANEXO 2 - MODELIZAÇÃO DE DY

Na modelização de DY seguiu-se a metodologia de Box e Jenkins caracterizada pelo prosseguimento de três fases: **Identificação** do processo estocástico subjacente à realização em causa, **Estimação** do modelo identificado, e **Análise diagnóstico** da estimação.

A2.1 - IDENTIFICAÇÃO

Visualizando o cronograma de DY (gráfico A2.1) constata-se que a série parece ter média estacionária. Este facto não é totalmente inesperado visto que, DY resulta já da diferenciação de primeira ordem do logaritmo do Rendimento Disponível a preços constantes de 1970.

GRAF. A2.1 - CRONOGRAMA DE DY



A identificação do modelo teórico correspondente a DY passa pela análise conjugada das respectivas função de autocorrelação (FAC) empírica e função de autocorrelação parcial (FACP) empírica. De seguida apresentam-se os resultados obtidos para estas funções.

FUNÇÃO DE AUTOCORRELAÇÃO DE DY

t	1	2	3	4	5	6
$\rho(t)$	0.413	0.167	0.150	-0.020	-0.011	0.209
	(2.15)	(0.75)	(0.66)	(-0.09)	(-0.05)	(0.91)

FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL DE DY

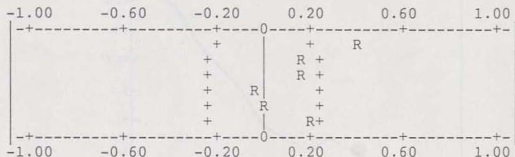
t	1	2	3	4	5	6
$\psi(t)$	0.413	-0.004	0.100	-0.141	0.073	0.222
	(2.15)	(0.02)	(0.52)	(-0.73)	(0.38)	(1.17)

Entre parêntesis encontra-se o valor da estatística t .

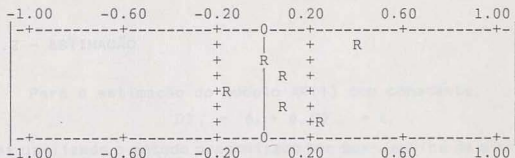
Dado que a série contém apenas 26 valores, as autocorrelações e autocorrelações parciais, só podem ser estimadas com uma certa fiabilidade para um número relativamente pequeno de lags (Chatfield(1984) aconselha um valor próximo de $N/4$) dificultando a identificação do modelo. Observando os gráficos da FAC e FACP empíricas nota-se que o comportamento destas funções não é muito típico. Ambas têm apenas os dois primeiros valores significativamente não nulos (com quocientes t superiores a dois), podendo espelhar tanto um AR(1), ou um MA(1) ou mesmo um ARMA(1,1). A função de densidade espectral (f.d.e.) com janela de Parzen, patente no gráfico A2.2, também não se ajusta inteiramente às formas conhecidas de processos teóricos. No entanto, parece estar

GRAF. A2.2 - FDE DF DY

AUTOCORRELATION FUNCTION OF THE SERIES $(1-B)^0 (1-B^0)^0$ DY

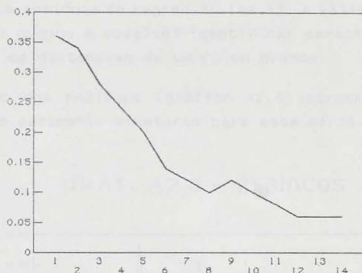


PARTIAL AUTOCORRELATION FUNCTION OF THE SERIES $(1-B)^0 (1-B^0)^0$ DY



mais próxima de um AR(1) com parâmetro positivo. Por estes motivos, e atendendo também ao princípio da parcimónia (sobre os parâmetros a estimar), optou-se por modelizar DY de acordo com um AR(1).

GRAF.A2.2 - FDE DE DY



A2.2 - ESTIMAÇÃO

Para a estimação do modelo AR(1) com constante,

$$DY_t = \phi_0 + \phi_1 DY_{t-1} + \epsilon_t ,$$

foi utilizado o método preconizado por Box-Jenkins de minimização dos quadrados dos resíduos com *backforecast*.

O modelo estimado obtido é igual a,

$$DY_t = 0.02612 + 0.42017 DY_{t-1} . \quad (2.496) \quad (2.294) \quad (A2.1)$$

O valor assumido pela estatística *t* indica que as estimativas dos parâmetros são significativamente não nulas. Por outro lado, como

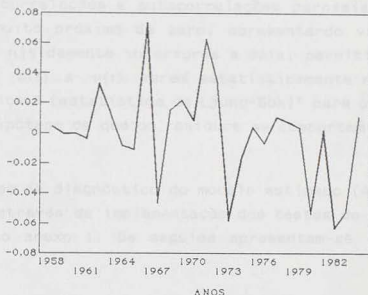
$|\hat{\phi}_1| < 1$ está assegurada a estacionariedade do processo.

A2.3 - ANÁLISE DE DIAGNOSTICO

A avaliação do modelo estimado passa pelo estudo do comportamento dos resíduos da regressão (A2.1). A validade do modelo é questionada quando é possível identificar características nos resíduos que os distanciam de um ruído branco.

O gráfico dos resíduos (gráfico A2.3) parece traduzir um comportamento puramente aleatório para esta série.

GRAF. A2.3 - RESÍDUOS



Passando à análise do Periodograma cumulativo (normalizado) dos resíduos, nota-se que todos os valores se encontram nitidamente dentro das bandas de confiança a 10%, estando muito próximos dos valores esperados para um ruído branco.

A FAC e FACP estimadas assumem os seguintes valores:

FUNÇÃO DE AUTOCORRELAÇÃO DOS RESÍDUOS

t	1	2	3	4	5	6
$\rho(t)$	-0.009 (-0.05)	-0.030 (-0.16)	0.161 (0.83)	-0.111 (-0.56)	-0.063 (-0.31)	0.236 (1.17)
Q	0.00	0.03	0.83	1.23	1.38	3.23

FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL DOS RESÍDUOS

t	1	2	3	4	5	6
$\psi(t)$	-0.009 (-0.05)	-0.030 (-0.16)	0.161 (0.84)	-0.113 (-0.59)	-0.054 (-0.28)	0.214 (1.11)

As autocorrelações e autocorrelações parciais dos resíduos situam-se muito próximo de zero, apresentando valores para o quociente t nitidamente inferiores a dois, permitindo aceitar a hipótese de $\rho(t)$ e $\psi(t)$ serem estatisticamente nulos. O valor da estatística Q (estatística de Ljung-Box)¹ para o lag 6 permite aceitar a hipótese de que os resíduos se comportam como um ruído branco.

A análise de diagnóstico do modelo estimado (A2.1) pode ser completada através da implementação dos testes de especificação inseridos no anexo 1. De seguida apresentam-se os resultados obtidos.

1 A estatística de Ljung-Box permite testar se os valores da FAC correspondem aos referentes a um ruído branco contra a alternativa de serem provenientes de um ARMA(p,q). Como tal a análise desta estatística tem maior significado na fase de avaliação da estimação. O valor de Q no último lag considerado, seja K , tem distribuição $\chi^2_{(K-q-p)}$.

AUTOCORRELAÇÃO

D70	GOD1	GOD2	GOD3	GOD4
-0.1173	0.0001	0.0010	0.0269	0.0392

HETEROESCEDASTICIDADE

ENG1	WHITE	BPAG1	BPAG2
0.0125	0.9456	0.4328	0.4398

NORMALIDADE

$\sqrt{b_1}$	b_2	BB	SW	SF
-0.1234	0.0406	0.0170	0.0852	0.0945

De acordo com os valores apresentados é possível concluir pela aceitação das hipóteses (tanto a um nível de significância de 5% como de 10%) de que as variáveis residuais não são correlacionadas, são homoscedásticas, e têm uma distribuição Normal. É possível ainda, através do teste de Jarque e Bera (vide anexo 1), aceitar a hipótese conjunta de independência, homoscedasticidade e normalidade dos ϵ_t ($t=1, \dots, 26$).

A análise de diagnóstico elaborada neste ponto permite concluir que a modelização apontada para DY é bastante satisfatória.

BIBLIOGRAFIA

- Abramovitch, L.; Singh, K. (1985). Edgeworth Corrected Pivotal Statistics and the Bootstrap. *Ann. Stat.*, **13**, 1, 116-132.
- Babu, G.J. (1989). Applications of Edgeworth Expansions to Bootstrap: A Review. *Data Analysis and Inference*, Neuchâtel, Yadolah Dodge, North-Holand.
- Babu, G.J.; Bose A. (1989a). Bootstrap Confidence Intervals. *Statistics & Probability Letters*, **7**, 2, 151-160.
- Babu, G.J.; Bose A. (1989b). *Accuracy of the Bootstrap Approximation*, Purdue University, Technical Report # 89-10.
- Babu, G.J.; Singh, K. (1983). Inference on Means Using the Bootstrap. *Ann. Stat.*, **11**, 3, 999-1003.
- Beran, R. (1982). Estimated Sampling Distributions: The Bootstrap and Competitors. *Ann. Stat.*, **10**, 3, 212-225.
- Beran, R. (1987). Prepivoting to Reduce Level Error of Confidence Sets. *Biometrika*, **74**, 3, 457-468.
- Beran, R. (1988). Prepivoting Test Statistics: a Bootstrap View of Asymptotic Refinements. *J.A.S.A.*, **83**, 403, 6878-697.
- Bhattacharya, R.N.; Ghosh, J.K. (1978). On the Validity of the Formal Edgeworth Expansion. *Ann. Stat.*, **6**, 2, 434-451.
- Bhattacharya, R.N.; Qumsiyeh, M. (1989). Second Order and Lp-Comparisons Between the Bootstrap and Empirical Edgeworth Expansion Methodologies. *Ann. Stat.*, **17**, 1, 160-169.
- Bickel, P.J.; Freedman, D.A. (1981). Some Asymptotic Theory for the Bootstrap. *Ann. Stat.*, **9**, 6, 1196-1217.
- Bose, A. (1988). Edgeworth Correction by Bootstrap in Autoregressions. *Ann. Stat.*, **16**, 4, 1709-1722.

- Box, G.E.P.; Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day.
- Breusch, T.S.; Pagan, A.R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47, 5, 1287-1294.
- Cartaxo, R.J.; Rosa, N.E.S. (1986). Séries Longas para as Contas Nacionais Portuguesas 1958-1985. *Banco de Portugal, Gabinete de Estudos, Documento de Trabalho* nº15.
- Chatfield, C. (1985). *The Analysis of Time Series: an Introduction*, third ed, Chapman and Hall.
- Davidson, A.C.; Hinkley, D.V.; Schechtman, E. (1986). Efficient Bootstrap Simulation. *Biometrika*, 73, 3, 555-566.
- Diccicio, T.J.; Romano, J.P. (1988). A Review of Bootstrap Confidence Intervals. *J.R.S.S., B*, 50, 3, 338-370.
- Diccicio, T.J.; Tibshirani, R. (1987). Bootstrap Confidence Intervals and Bootstrap Approximations. *J.A.S.A.*, 82, 397, 163-170.
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.*, 7, 1, 1-26.
- Efron, B. (1980). Computer Intensive Methods in Statistics. *Some Recent Advances in Statistics*, Lisboa, J. Tiago de Oliveira e Benjamim Epstein, Academia das Ciências de Lisboa.
- Efron, B. (1981). Nonparametric Standard Errors and Confidence Intervals. *The Canadian Journal of Statistics*, 9, 2, 139-172.
- Efron, B. (1982a). The Jackknife, the Bootstrap and Other Resampling Plans. *Society for Industrial and Applied Mathematics*.
- Efron, B. (1982b). Transformation Theory: How Normal is a Family of Distributions? *Ann. Stat.*, 10, 323-339.
- Efron, B. (1985). Bootstrap Confidence Intervals for a Class of Parametric Problems. *Biometrika*, 72, 1, 45-58.

- Efron, B. (1987). Better Bootstrap Confidence Intervals. *J.A.S.A.*, 82, 397, 171-200.
- Efron, B.; Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, 37, 1, 36-48.
- Efron, B.; Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1, 1, 54-77.
- Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- Freedman, D.A. (1981). Bootstrapping Regression Models. *Ann. Stat.*, 9, 6, 1218-1228.
- Freedman, D.A. (1984). On Bootstrapping Two-Stages Least-Squares Estimates in Stationary Linear Models. *Ann. Stat.*, 12, 3, 827-842.
- Freedman, D.; Peters, S.C. (1984). Bootstrapping a Regression Equation: Some Empirical Results. *J.A.S.A.*, 79, 385, 97-106.
- Ghosh, M.; Parr, W.C.; Singh, K.; Babu, G.J. (1984). A Note on Bootstrapping the Sample Median. *Ann. Stat.*, 12, 3, 1130-1135.
- Godfrey, L.G. (1978). Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables. *Econometrica*, 46, 6, 1293-1302.
- Goldberger (1964). *Econometric Theory*, Wiley.
- Griffiths, W.E.; Surekha, K. (1986). A Monte Carlo Evaluation of the Power of Some Tests for Homoscedasticity. *Journal of Econometrics*, 31, 219-231.
- Hall, P. (1983). Inverting an Edgeworth Expansion. *Ann. Stat.*, 11, 2, 569-576.
- Hall, P. (1986). On the Bootstrap and Confidence Intervals. *Ann. Stat.*, 14, 4, 1431-1452.

- Hall, P. (1987). On the Bootstrap and Continuity Correction. *J.R.S.S.*, 49, 1, 82-89.
- Hall, P. (1988). Theoretical Comparison of Bootstrap Confidence Intervals. *Ann. Stat.*, 16, 3, 927-985.
- Hall, P.; Thomas, J.; DiCiccio; Romano, J.P. (1989). On Smoothing and The Bootstrap. *Ann. Stat.*, 17, 2, 692-704.
- Harter, L. (1961). Expected Values of Normal Order Statistics, *Biometrika*, 48, 151-165.
- Hinkley, D.V. (1988), (com discussão). Bootstrap Methods. *J.R.S.S.*, B, 50, 3, 321-337, 355-370.
- Jarque, M.C.; Bera, A.K. (1980). Efficient Tests for Normality, Homoscedasticity, and Serial Independence of Regression Residuals. *Economics Letters*, 6, 255-259.
- Judge, G.C.; Griffiths, W.E.; Luetkepohl, H.; Lee, T. (1982). *Introduction to the Theory and Practice of Econometrics*, New York, Wiley.
- Judge, G.C.; Griffiths, W.E.; Luetkepohl, H.; Lee, T. (1985). *The theory and Practice of Econometrics*, 2nd ed., New York, Wiley.
- Kiefer, N.M.; Salmon, M. (1983). Testing Normality in Econometric Models. *Economics Letters*, 11, 123-128.
- Liu, R.Y.; Singh, K. (1987). On a Partial Correction by The Bootstrap. *Ann. Stat.*, 15, 4, 1713-1718.
- Loh, W.Y. (1987). Calibrating Confidence Coefficients. *J.A.S.A.*, 82, 397, 155-162.
- Murteira, B.J.F (1988). *Intervalos de Confiança "Bootstrap". Aplicação ao Índice de Gini*. Centro de Matemática Aplicada à Previsão e Decisão Económica, Instituto Superior de Economia, Documento de Trabalho nº48.
- Pindyck, R.S.; Rubinfeld D.L. (1976). *Econometric Models and Econometric Forecasts*, New York, McGraw-Hill Inc.

- Prescott, D.M.; Stengos, T. (1987). Bootstrapping Confidence Intervals: An Application to Forecasting the Supply of Pork. *American Journal of Agricultural Economics*, 69, 2, 266-273.
- Proença, I.M.D. (1988). *Estimativas "Jackknife" e "Bootstrap" para o Enviesamento e Desvio Padrão do Índice de Gini*. Centro de Matemática Aplicada à Previsão e Decisão Económica, Instituto Superior de Economia, Documento de Trabalho nº67.
- Rayner, R.K. (1988). Some Asymptotic Theory for the Bootstrap in Econometric Models. *Economics Letters*, 26, 43-47.
- Rubin, D.B. (1981). The Bayesian Bootstrap. *Ann. Stat.*, 9, 1, 130-134.
- Santos Silva, J.M.C. (1988). *Função Consumo: Alguns Desenvolvidos Recentes e Análise do Caso Português*, Instituto Superior de Economia, Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Métodos Matemáticos para Economia e Gestão.
- Schenker, N. (1985). Qualms About Bootstrap Confidence Intervals. *J.A.S.A.*, 80, 390, 360-361.
- Shao, J.(1988a). On Resampling Methods for Variance and Bias Estimation in Linear Models. *Ann. Stat.*, 16, 3, 986-1008.
- Shao, J.(1988b). Bootstrap Variance and Bias Estimation in Linear Models. *The Canadian Journal of Statistics*, 16, 4, 371-382.
- Shapiro, S.S.; Francia, R.S. (1972). An approximate Analysis of Variance Test for Normality. *J.A.S.A.*, 67, 215-216.
- Shapiro, S.S.; Wilk, M.B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52, 3 e 4, 591-611.
- Silverman, B.W.; Young, G.A. (1987). The Bootstrap: To Smooth or Not Smooth? *Biometrika*, 74, 3, 469-479.
- Singh, K. (1981). On the Asymptotic Accuracy of Efron's Bootstrap. *Ann. Stat.*, 9, 6, 1187-1195.

- Tibshirani, R. (1988). Variance Stabilization and The Bootstrap. *Biometrika*, 75, 3, 433-444.
- Veall, M.R. (1987a). Bootstrapping forecast uncertainty: A Monte Carlo Analysis. *Time Series and Econometric Modelling*, Dordrecht, I.B. MacNeil; G.J. Umphrey, 203-212.
- Veall, M.R. (1987b). Bootstrapping the Probability distribution of peak electricity demand. *International Economic Review*, 28, 203-212.
- Weisberg, S. (1980). *Applied Linear Regression*, New York, Wiley.
- White, H.(1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48, 4, 817-838.
- White, H.; MacDonald, G.M. (1980). Some Large-Sample Tests for Nonnormality in the Linear Regression Model. *J.A.S.A.*, 75, 369, 16-31.
- Wu, C.F.J. (1986). Jackknife, Bootstrap, and Other Resampling Methods in Regression Analysis. *Ann. Stat.*, 14, 4, 1261-1350.
- Young, G.A. (1988). A Note on Bootstrapping the Correlation Coefficient. *Biometrika*, 75, 2, 370-373.

LISTA DAS ABREVIATURAS USADAS:

Ann. Stat. - ANNALS OF STATISTICS

J.A.S.A. - JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

J.R.S.S. - JOURNAL OF ROYAL STATISTICAL SOCIETY